Quantitative functional brain imaging with Positron Emission Tomography and irreversible tracers

Veronese, Mattia

Awarding institution:
King's College London
University of Padua

Departiment of Information Engineering

Padova University, Padova, Italy

Ph.D. School on Information Engineering

Section: *Bioengineering* - Cycle: *XXV*

A non compartmental method for functional quantitative imaging with Positron Emission Tomography and irreversible tracers

Headmaster of the school: Prof. Matteo Bertocco

Supervisor: Prof. Alessandra Bertoldo

PhD candidate: Mattia Veronese

To myself, a selfish but very satisfying gift.

A non compartmental method for functional quantitative imaging with Positron Emission Tomography and irreversible tracers

ABSTRACT

In dynamic Positron Emission Tomography (PET) studies the term *spectral analysis* indicates a time-invariant single input/single output model, used for the data quantification [Cunningham and Jones, 1993]. Despite the name and its common use in the engineering field, SA does not indicate an analysis in the frequency domain but, instead, it represents a method from which the radioactivity concentration measured with PET can be related to the underlying physiological processes of the investigated system. SA is so-called, because it provides a spectrum of the kinetic components from which it is possible to derive a large variety of physiological parameters, depending on the characteristics of the analyzed tracers.

In the last years SA has been widely used with a large number of PET tracers to study brain and non brain tissues, demonstrating to be a very flexible method. Differently from the most used PET quantification approaches, like the compartmental modelling [Godfrey, 1982] or the graphical methods [Patlak, 1983; Logan *et al.*, 1990], SA can be applied to homogeneous as well as to heterogeneous kinetic tissues without any specific compartmental model assumption. This characteristic makes it a high informative investigative tool especially for the analysis of novel PET tracers.

The most critical aspect of SA is related to its sensitivity to the presence of noise in the data. This characteristic makes SA not properly indicated for the application to low signal-to-noise ratio (SNR) data [Turkheimer *et al.*, 1994]. During the past several years, several solutions have been introduced to improve the robustness of SA in the presence of noise. The most famous example is represented by rank-shaping spectral analysis (RS) [Turkheimer *et al.*, 2003]. However, even if RS has been shown to be a precise and accurate quantification method, its applicability is limited to tracers with reversible uptake. This is a severe restriction if we consider that one of the most used PET tracer for clinical research, 18F-Fluorodeoxyglucose ([^{18}F]FDG), is irreversible.

In this work we present SAIF, (Spectral Analysis with Iterative Filter), a SA-based method for the quantification of PET data investigated with irreversible-uptake tracers. SAIF has been designed in order to maintain the main advantages of SA but providing a superior robustness to measurement noise. The final aim was to create a reliable and flexible PET quantification tool, offering a valid alternative to standard methodologies for functional quantitative imaging with PET and irreversible tracers.

The organization of this thesis is as follows: Chapter 1 offers a brief introduction to PET technique and its quantification methods. A comparison between compartmental modelling approaches and graphical methods is also presented, in order to provide the operative context in which SA is located. Chapter 2 contains the mathematical formalization of the SA model. Standard and filtered SA versions are presented with particular attention to novelty elements introduced by SAIF. In Chapter 3 and

Chapter 4, SAIF will be tested with brain and non brain PET data. Several datasets obtained by using different PET tracers are considered. As an example for brain tissue quantification, SAIF application to L-[1-^{11}C]Leucine and [^{11}C]SCH442416 data is presented. For non brain tissues, instead, analysis of three datasets is reported: 1) [^{18}F]FDG PET studies applied to skeletal leg muscle, 2) [^{18}F]FLT PET studies applied to breast cancer patients and 3) [^{18}F]FDG PET studies applied to normal control and acute lung injury patients. For each dataset SAIF results are compared with those provided by already validated methods and used in the literature as reference for the quantification. This analysis allows to compare SAIF performances with those offered by the current state of the art. Chapter 5 investigates the conditioning of the kinetic heterogeneity to PET quantification. The relationship between this problem, the spatial resolution of the imaging technique and the noise level of the data is also considered. This aspect is a critical point for PET quantification because when it is not taken into account it can lead to heavily biased results. Particular attention is given to how SAIF addresses this issue. In Chapter 6 we present SAKE, a software application developed in-house which implements the major SA algorithms. SAKE manages the whole process of PET quantification: from data preprocessing to the result analysis. No other program or additional tool is required. Chapter 7 discusses the most relevant criticalities of the SA approach and of SAIF method in particular. Considerable attention is given to the definition of the setting algorithm as well as to the model assumptions used by SAIF to describe the data. In Chapter 8 an overall discussion is presented with a conclusive sum-

mary about strengths and weakness of SAIF method. The appendix of the thesis is dedicated to the some additional works, not directly related to the main argument of this PhD project, but of interest for the PET field. This research concerns 1) the development of voxelwise quantification methods for [^{11}C](R)Rolipram PET data, 2)the use of non linear mixed effects modelling for plasma metabolite correction, and 3) the evaluation of the sensitivity of PET receptor occupancy studies to the experimental design.

Un metodo non compartimentale per l'imaging funzionale quantitativo con Tomografia ad Emissione di Positroni e traccianti irreversibili

Sommario

Negli studi dinamici di Tomografia ad Emissione di Positroni (dall'inglese Positron Emission Tomography, PET) il termine analisi spettrale (SA) indica un modello tempo-invariante singolo-ingresso/singola-uscita per la descrizione temporale dei dati acquisti dall'esame [Cunningham and Jones, 1993]. Nonostante l'accezione con cui viene comunemente utilizzata all'interno del contesto ingegneristico, SA non fa riferimento ad un'analisi nel dominio delle frequenze quanto, piuttosto, rappresenta un metodo attraverso cui la radioattività misurata durante l'esame PET viene messa in relazione con i processi fisiologici del sistema investigato. Il metodo SA è così chiamato perchè fornisce lo spettro cinetico dell'attità del tracciante nei tessuti, a partire dal quale è possibile derivare una grande varietà di parametri fisiologici relativi al sistema in esame.

Come riportato negli studi di Tadokoro (1993), Fujiwara (1996), Richardson (1996), Bertoldo (1998), Hinz (2008), Brooks (2008) e Myers (2012), negli ultimi anni SA è stata ampiamente utilizzata per lo studio dei tessuti cerebrali e non, dimostrando di essere uno strumento adattabile ad una grande varietà di traccianti e processi. Allo stesso tempo SA richiede minime assunzioni per la sua applicazione: differentemente da molti approcci di quantificazione di immagini PET, come la modellistica compartimentale [Godfrey, 1983] o i metodi grafici [Patlak, 1983; Logan *et al.*,

1990], SA può essere applicata ai tessuti omogenei così come ai tessuti eterogenei senza la necessità di ulteriori assunzioni. Questa caratteristica fa della SA uno strumento investigativo altamente informativo specialmente per l'analisi di nuovi traccianti PET in cui le conoscenze a priori risultano limitate.

La maggiore criticità della tecnica SA è legata alla sensitività della metodologia alla presenza di rumore nei dati, caratteristica che ne limita l'applicabilità in caso di basso rapporto segnale rumore (SNR) [Turkheimer *et al.*, 1994]. Durante gli anni svariate alternative sono state introdotte per migliorare la robustezza della tecnica alla presenza di rumore. La soluzione di maggior successo è rappresentata dalla rank-shaping spectral analysis (RS) [Turkheimer *et al.*, 2003]. Tuttavia, sebbene sia stato dimostrato come RS rappresenti un preciso e accurato metodo di quantificazione, l'applicabilità della tecnica resta limitata ai soli traccianti con cinetica reversibile. Tale peculiarità costituisce una severa restrizione all'uso della RS se si considera che molti traccianti non rientrano in questa categoria. L'esempio più importante è rappresentato dal [^{18}F]FDG (Fluorodeoxyglucose), il tracciante PET più impiegato al mondo nella ricerca clinica e pratica medica.

In questo lavoro viene presentata SAIF (Spectral Analysis with Iterative Filter), un versione filtrata della spectral analysis per la quantificazioni di dati PET ottenuti a partire da traccianti con cinetica irreversibile. SAIF è stata sviluppata per mantenere i principali vantaggi della SA ma allo stesso tempo si propone di offrire una superiore robustezza all'errore di misura.

Il contenuto della presente tesi si articola come segue: il capitolo 1 offre una breve introduzione alla tecnica PET e ai principali metodi di quantificazione. Un confronto tra metodi compartimentali e metodi grafici viene presentato in modo da definire le caratteristiche del contesto in cui la tecnica SA si colloca. Il capitolo 2 contiene la formalizzazione matematica della SA. Il metodo standard e le relative versioni filtrate vengono presentante, prestando particolare attenzione agli elementi di novità introdotti dalla tecnica SAIF. Nei capitoli 3 e 4 SAIF viene testata in diversi casi di studio, rispettivamente per esami PET cerebrali e non. Come esempi per la quantificazione dei tessuti cerebrali si presenteranno i casi di applicazione della SAIF ai traccianti L-[1-^{11}C]Leucina e [^{11}C]SCH442416. Relativamente ai tessuti non cerebrali, invece, si prenderanno in considerazione tre casi di interesse: 1) quantificazione di immagini PET ottenuti con tracciante [^{18}F]FDG per lo studio del muscolo scheletrico; 2) quantificazione di immagini PET ottenuti con tracciante [^{18}F]FLT in pazienti oncologici con tumori al seno; 3) quantificazione di immagini PET ottenuti con tracciante [^{18}F]FDG per lo studio dei tessuti polmonari in soggetti sani e in pazienti affetti da sindrome polmonare acuta. Per ogni dataset i risultati ottenuti con il metodo SAIF saranno confrontati con quelli ottenuti da metodi di quantificazione già validati e indicati dalla letteratura come standard di misura. Tali analisi permetteranno di capire le performance della SAIF rispetto al corrente stato dell'arte metodologico. Il capitolo 5 approfondisce il condizionamento della eterogeneità cinetica dei sistemi biologici sulla quantificazione delle immagini PET, analizzando la dipendenza del problema rispetto alla risoluzione spaziale delle im-

magini e al rumore dei dati. Particolare attenzione verrà riservata a come la metodologia SAIF si approccia al problema. Nel capitolo 6 verrà introdotto SAKE, un'applicazione software sviluppata per implementare i maggiori algoritmi di spectral analysis. Il capitolo 7 propone un'analisi delle criticità degli approcci SA in generale e SAIF in particolare. Verrà discusso come definire il setting applicativo degli algoritmi così come gestire le assunzioni cinetiche dei vari approcci modellistici. In fine verranno presentate le considerazioni complessive dell'intero lavoro, proponendo un riassunto dei punti di forza e debolezza del metodo SAIF.
L'appendice della tesi è dedicata ad alcuni lavori addizionali non direttamente collegati all'argomento principale di questa tesi di dottorato, ma di interesse per l'ambito PET. Le ricerche si riferiscono 1) allo sviluppo di metodi voxel-wise per la quantificazione di dati PET ottenuti con tracciante [^{11}C](R)Rolipram, 2) all'uso del metodo non linear mixed effects modelling per la correzione dei metaboliti plasmatici, e 3) alla valutazione della sensitività degli studi recettoriali PET al protocollo sperimentale.

Contents

If the facts don't fit the theory, change the facts.

Albert Einstein

1

Introduction to quantitative PET

POSITRON EMISSION TOMOGRAPHY (PET) is a nuclear medicine imaging technique that produces *in vivo* three-dimensional images of functional dynamic processes in the body. The system detects pairs of gamma rays emitted indirectly by a positron-emitting radionuclide (tracer), which is introduced into the body on a biologically active molecule. Thus, depending on the characteristics of the injected compound, it is possible to derive a large variety of physiological parameters such as blood flow, glucose metabolism or neuroreceptor binding. In contrast to structural imaging techniques, like X-ray or MRI, PET can be used to examine dynamic changes in selected functional processes showing not only where these changes takes place but also measuring their magnitude. In order to relate the radioactivity concentration measured with PET to the underlying physiological or biochemical processes, the application of mathematical models to describe tracer kinetics within an area of interest is

necessary. For this purpose, several solutions are available.
In clinical practise PET quantification is routinely performed by using the Standard Uptake Value (SUV), a semi-quantitative index in use since the 1980s (Kubota, 1985). Its computation is very simple since it requires only the PET measure at a pre-fixed sample time and the injected dose normalized by some anthropometric characteristics of the subject (generally the body weight or the body surface area) (Kim et al., 1994; Thie, 2004). SUV is characterised by general applicability and no additional assumption concerning reversibility/irreversibility of the tracer and the system under study is necessary. However its use is controversial because SUV computation is affected by many technical and physiological factors leading to biased results (Boellaard et al., 2004; Hamberg et al., 1994; Huang, 2000).
An alternative to SUV is represented by the tissue-to-plasma ratio (RATIO) method (Lehtio et al., 2003; Mitkovski et al., 2005). As indicated by the name, this index is computed as the ratio between the activity of the tracers measured in the tissue and the activity of the tracers measured in the plasma pool within a pre-fixed time window. Fundamental requirement for RATIO definition is the presence of kinetic equilibrium in the system of interest. As with the SUV, the RATIO is a very simple and fast index to compute and it is generally applicable to most of the PET tracers. However, both SUV and RATIO do not return any information about the kinetics of the investigated system and therefore their application in dynamic PET studies is poorly informative. For this reason SUV and RATIO are more commonly applied as diagnostic indexes rather than quantitative PET methods for clinical research.
The most used approach for quantification of dynamic PET studies is represented by compartmental modelling (Godfrey, 1982). This method is based on a first-order differential description of the main physiological processes in which the tracer is involved. Once the best compartmental structure describing the behaviour of the tracer in the tissues has been defined, it is possible to estimate the system's micro and macro parameters. This can be achieved at the Region of Interest (ROI) level, i.e. generating

a tissue Time Activity Curve (TAC) of the average activity concentration within a specific tissue area, or at the voxel level, i.e. deriving the tissue TAC from one single image unit. The ROI level approach is commonly used to quantify PET images by using nonlinear estimators since ROI TACs are characterized by a good signal-to-noise ratio (SNR). However, the use of ROI TACs causes a loss of spatial resolution since it does not allow the quantification of the physiological information with the same spatial detail as the acquired PET data. Voxel level analysis allows one to avoid the operator-dependency of manually delineated ROIs and produces parametric maps having the same spatial resolution as the original PET image. However, TACs derived from single voxels are characterized by a low SNR. This renders the use of nonlinear estimators difficult and unwieldy because of their computational cost. Thus, more robust and faster estimation algorithms are necessary in order to quantify the physiological information at the voxel level within a reasonable computational time.

To improve robustness and/or computational efficiency of voxel-wise quantification level via compartmental modelling several solutions have been proposed in literature. The most used are represented by the generalized linear least square (GLLS) (Feng et al., 1996), the basis function methods (BFMs) (Gunn et al., 1997; Hong and Fryer, 2010; Rizzo et al., 2012a), Global Two Stage (GTS) (Tomasi et al., 2009a, 2011a), and Bayesian estimation (Alpert and Yuan, 2009; Rizzo et al., 2012b). However, even if these methodologies have been shown to provide some improvements compared to the standard compartmental modelling approach, each one presents its own weaknesses which limit the applicability of these methodologies.

A valid alternative to compartmental modelling is represented by Patlak (Patlak and Blasberg, 1985) and Logan (Logan et al., 1990) graphical methods. Both of them have been widely used for PET quantification, because of their simplicity and model independence. In fact, they do not need any *a priori* definition of the model structure, requiring only the information on the reversibility or irreversibility of the tracer kinetics. Be-

ing easy to implement, they have been routinely used for the generation of parametric maps. Nevertheless, they are affected by several limitations. First of all both methods allow the estimation of a unique macroparameter (i.e. the tracer net trapping uptake for Patlak and the tracer distribution volume for Logan) without fully describing the tracer kinetics. At the same time they do not account for the vascular component. Moreover, due to the noise in the data, they may lead to biased parameter estimates, especially when applied voxel-wise (Slifstein and Laruelle, 2001). Another approach for PET data analysis is represented by Spectral Analysis (SA) (Cunningham and Jones, 1993). In dynamic PET studies the term *spectral analysis* indicates a single-input/single-output model used for the data quantification. Thus, it is important to note that SA is not an analysis in the frequency domain but it is so-called because it provides a spectrum of the kinetic components from which it is possible to derive a selected set of physiological parameters independently from the specific model configuration. Even if in PET community SA belongs to the group of data-driven methods, for its application it requires the fulfilment of some assumptions about the tracer kinetics. First, the SA model is defined only for single input systems (Schmidt, 1999). Moreover SA is not applicable to cyclic systems, a condition which is very uncommon between PET tracers (Schmidt, 1999). On the other hand, SA allows to obtain the complete description of tracer kinetics as well as the number and the type of compartments necessary for the data modelling (Cunningham and Jones, 1993; Turkheimer et al., 1994). This information can be very useful for the specification of the kinetic model of a novel PET tracer. Compared to graphical methods and compartmental modelling approaches, these SA features represent a good trade-off between the information obtainable by the method and the requirements for its application.

Figure 1.0.1 reports a schematic summary of the major PET quantification methods organized by considering for each method the information returned as function of its application requirements. Within this diagram it results clearly evident a diagonal distribution of the methodologies.

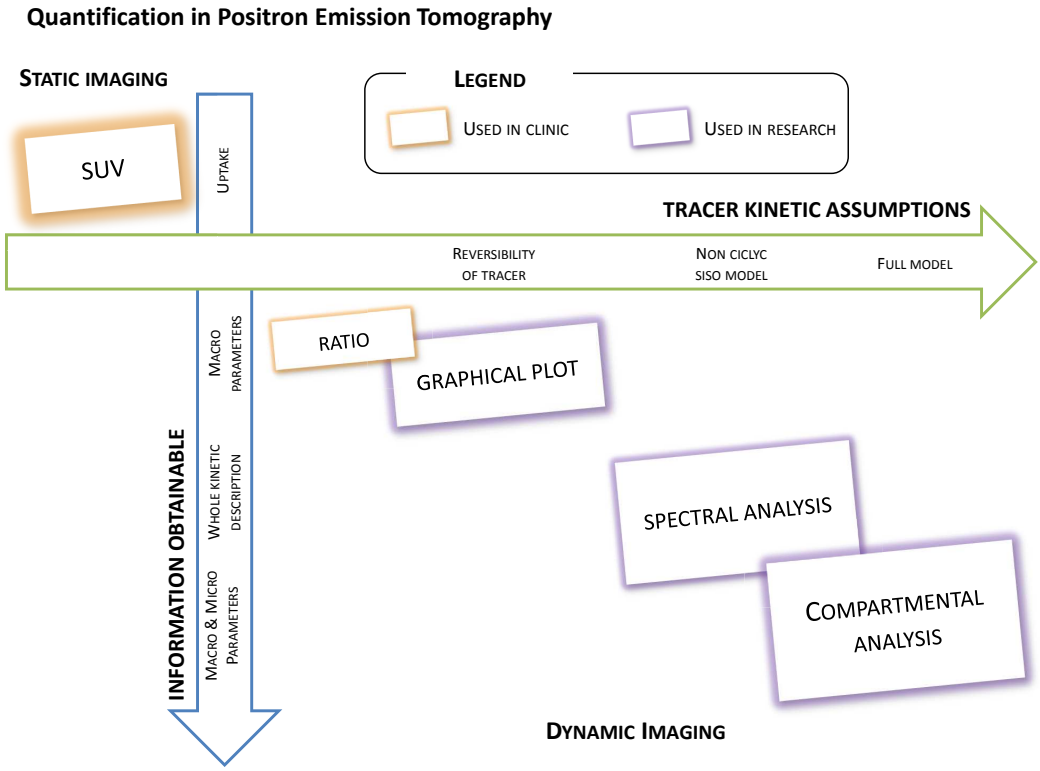

Figure 1.0.1: Quantification in Positron Emission Tomography. The figure shows a schematic representation of the main quantification methods used in PET. The spatial position of each method in the diagram indicates the information obtainable by the method at the cost of the assumptions necessary for its application. Orange and purple colours indicate respectively the methods used in clinical practise and in research. Static and dynamic imaging are separately reported.

This indicates that it is possible to choose a more informative method only at the cost of additional assumptions about the investigated system.

Aerodynamically, the bumble bee shouldn't be able to fly, but the bumble bee doesn't know it so it goes on flying anyway.

Mary Kay Ash

2

A non compartmental method for quantitative PET imaging: Spectral Analysis

2.1 The input/output Spectral Analysis model

2.1.1 Mathematical Formalization

In spectral analysis, the concentration of radioactivity in the tissue at time t, $C_{tiss}(t)$, is modelled as a convolution of the plasma time-activity curve, $C_p(t)$, with the sum of $M+1$ distinct exponential terms as

$$C_{tiss}(t) = \sum_{j=0}^{M} C_p(t) \otimes \alpha_j \cdot e^{-\beta_j t} \tag{2.1}$$

where α_j and β_j ($\beta_1 < \beta_2 < \cdots < \beta_M$) are assumed to be real-valued and nonnegative. This constraint derives from the assumption that SA is modelling a first order compartmental system with a single arterial input. The upper limit, $M + 1$, represents the maximum numbers of terms to be included in the model and this is, in general, set to a large number to be established, usually 100. The values of β_j are predetermined and fixed in order to cover an appropriate spectral range. For the studies involving short lived positron emitting isotopes this range needs to extend to the slowest possible event of the tracer in the tissue up to a value appropriate to transient phenomena (e.g. the passage of activity through the tissue vasculature). The values of α_j are estimated from the blood and tissue time activity curves by a nonnegative least squares (NNLS) procedure. In practice, only a few components with $\alpha_j > 0$ are detected, originating what is called the kinetic spectrum of the tracer in the tissues (Fig. 2.1.1A). The estimated spectral components assume different meaning depending on the position of the beta grid where they are located. For example, the corresponding term for $\lim \beta_j \to \infty$ (i.e., components with β_j very large), become proportional to $C_p(t)$, and can be seen as "high-frequency" components. In the same way the corresponding term with $\beta_j = 0$ or near zero, becomes proportional to $\int C_p(\tau)d\tau$ and can be viewed as the "low-frequency" component, i.e. accounting for trapping of the tracer. Components with intermediate values of β_j ("equilibrating components") reflect tissue compartments that exchange material directly or indirectly with the plasma with their number corresponding to the number of identifiable tissue compartments within the region of interest. In light of these particular features, it is very common to define the SA model equation explicitly showing trapping in the following way

$$C_{tiss}(t) = \alpha_0 \cdot \int_0^t C_(\tau)d\tau + \sum_{j=1}^{M} C_p(t) \otimes \alpha_j \cdot e^{-\beta_j t} \tag{2.2}$$

where $\beta_j > 0$, $j = 1, 2, \ldots, M$.
The spectral analysis model has general applicability. However the func-

tional base $C_p(t) \otimes^{(-\beta_j t)}$ is an overcompleted base of the space of interest[1] (Schmidt and Turkheimer, 2002). This r results in two main problems: first of all the error properties of the estimates as well as the influence of the error on the estimated components are difficult to estimate and control (Turkheimer et al., 1998); secondly to identify an unique and sparse solution the coefficients α_j must be constrained to be positive. Unfortunately not all compartmental models satisfy these assumptions (Schmidt, 1999). Reference region models, for example, may result in negative α_j and therefore SA is not applicable (Gunn et al., 2001). On the other hand, all the models with plasma or blood input functions and noncyclic structures result in positive α_j values (Schmidt, 1999), allowing the applicability of SA. Most of the kinetic models used with many PET tracers met these conditions and therefore they do not represent any limitation for spectral analysis utilization (Schmidt, 1999).

2.1.2 Estimation of the parameters of interest

The main purpose of SA application to dynamic PET data is the quantitative characterization of the tracer kinetics within the target tissues. This is possible by linking the estimated spectral components, i.e. the estimated α_j and β_j, with a selection of parameters of interest that do not depend on a specific model representation. As demonstrated by Gunn and colleagues (2001) when a particular system meets the conditions to be modelled with spectral analysis[2] the unique idenfiability of some macroparameters of interest is guaranteed. These parameters coincide with the influx rate constant (K_1, $ml/cm^3/min$), the net uptake of the tracer in the

[1] In functional analysis, an orthogonal basis is any generalization of an Hilbert basis obtained from the multiplication between a orthonormal basis using multiplication by nonzero scalars. An orthonormal basis for a Hilbert space H is an orthonormal set of vectors with the property that every vector in H can be written as an infinite linear combination of the vectors in the basis.

[2] The SA applicability is limited to those systems in which the state transition matrix is negative semidefinite (Schmidt, 1999). This condition is met by all single-input/single-output noncyclic systems (Schmidt, 1999). The subset of cyclic systems in which the product of the rate constants is the same regardless of directions for every cycle respects the constraints to be modelled with SA (Godfrey, 1982).

tissues (K_i, $ml/cm^3/min$) and the volume of distribution (V_T, ml/cm^3)[3]. The last two elements cannot be estimated simultaneously, because they respectively depend on the irreversibility or reversibility of the tracer kinetic.

The relationship between the parameter values and estimated spectrum is as follows. The transport of tracer from plasma to tissue, K_1, coincides with the sum of components' amplitudes, i.e.

$$K_1 = \sum_{j=0}^{M} \alpha_j \tag{2.3}$$

In case of irreversible tracers, K_i can be derived by the limit of the SA model impulsive response function ($IRF(t)$) for $t \to \infty$, which is also equal to the amplitude of the estimated component corresponding to $\beta_j = 0$:

$$K_i = \lim_{t\to\infty} IRF(t) = \lim_{t\to\infty} \sum_{j=0}^{M} \{\alpha_j \cdot e^{-\beta_j t}\} = \alpha_0 \tag{2.4}$$

In case of reversible tracers, instead, V_T can be computed from the integral of $IRF(t)$ as:

$$V_T = \int_0^{\infty} IRF(\tau)d\tau = \sum_{j=1}^{M} \frac{\alpha_j}{\beta_j} \tag{2.5}$$

In addition to these parameters, if the measurement equation for the total radioactivity measured by the PET scanner takes into account the tracer contribution in both blood and tissues, it is also possible to derive the blood volume (V_b, unitless). Generally this corresponds to the case in which

$$C_{measured}(t) = (1 - V_b) \cdot C_{tiss}(t) + V_b \cdot C_b(t) \tag{2.6}$$

where $C_{measured}(t)$ represents the total activity measured by the scanner within a specified volume of observation, $C_{tiss}(t)$ represents the tissue

[3]The measurement units here reported for K_i, K_1 and V_T follow the guidelines of PET modeller consensus (Innis et al., 2007).

kinetic activity and $C_b(t)$ the blood tracer activity. The use of this measurement equation (Eq. 2.6) in combination with the SA model (Eq. 2.1) allows SA to fully characterized the whole time-course of tracer kinetics starting from the activity measured by the scanner as

$$C_{measured}(t) = (1 - V_b) \cdot \{\sum_{j=0}^{M} C_p(t) \otimes \alpha_j \cdot e^{-\beta_j t}\} + V_b \cdot C_b(t) \qquad (2.7)$$

From SA application it is also possible to characterize how the SA mo-

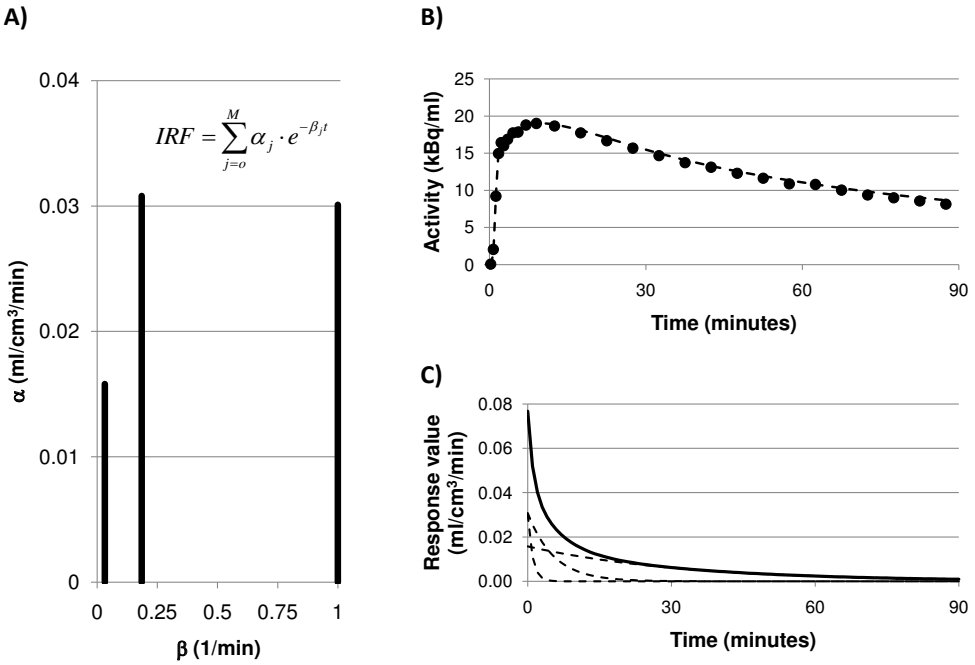

Figure 2.1.1: Spectral Analysis output. Panel A shows an example of a possible SA spectrum. Panel B shows the relative SA model fit (dashed line) to the measured data (black dots). Panel C reports the corresponding IRF time-activity (solid line) and the single component time-courses (dashed lines). IRF is computed from the SA model (Eq. 2.1) by assuming unit impulsive. Thus it results that $IRF(t) = \sum_{j=0}^{M} \alpha_j \cdot e^{-\beta_j t}$.

del-prediction fits to the measured data (Fig. 2.1.1B). This visual analysis

provides a first qualitative evaluation about the reliability of the SA data description. In addition, for each estimated component, it is also possible to compute the correspondent time-course (Fig. 2.1.1C), breaking down in single kinetic elements the total tracer activity.
From the analysis of the SA estimated spectrum, however, it is not possible to compute the micro-parameters of the system, unless a fully characterization of the compartmental model describing the system is known in advance. In fact, from the indistinguibility theorem (Gunn et al., 2001) any two plasma input models with a total number of N tissue compartments are not distinguishable. The lack of a unique bidirectional relationship between SA spectra and compartmental models, prevents the identification of a unique compartmental model given a specific spectrum and thus a unique set of system micro-parameters.
On the contrary, given a compartmental model which fulfills the SA requirements (i.e. noncyclic and with single arterial input), there is a fully-described relationship between the spectral components and the model configuration.

As example we can considered a two tissue compartmental model with trapping (Fig. 2.1.2A). This represents a very interesting case as it is applied for describing several tracer brain kinetics like [^{18}F]FDG (Phelps et al., 1979; Sokoloff et al., 1977) and L-[1-^{11}C]Leucine (Schmidt et al., 2005). With some renumbering the equations of the model are

$$\begin{cases} \frac{C_1(t)}{dt} &= K_1 C_p(t) - (k_2 + k_3) C_1(t) \\ \frac{C_2(t)}{dt} &= k_3 C_1(t) \\ y(t) &= C_1(t) + C_2(t) \end{cases} \tag{2.8}$$

corresponding to the explicit solution

$$y(t) = \frac{K_1 k_3}{k_2 + k_3} \int_0^t C_p(\tau) d\tau + \frac{K_1 k_2}{k_2 + k_3} \int_0^t C_p(\tau) \cdot e^{-(k_2+k_3)(t-\tau)} d\tau \tag{2.9}$$

In this case the spectrum provided by the SA consists of two components, a trapping component and an equilibrating component (Fig. 2.1.2B). The

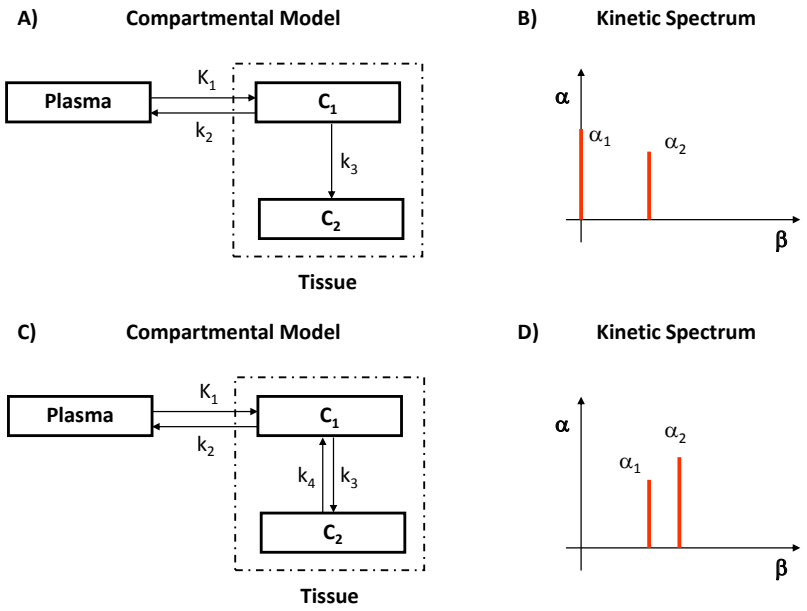

Figure 2.1.2: Two-tissue compartmental model with trapping (Panel A) and its relative spectrum (Panel B). Two-tissue compartmental model (Panel C) and its relative spectrum (Panel D).

associated SA equation is

$$y(t) = \alpha_0 \int_0^t C_p(\tau)d\tau + \alpha_1 \int_0^t C_p(\tau) \cdot e^{-\beta_1(t-\tau)}d\tau \tag{2.10}$$

where

$$\begin{cases} \frac{K_1 k_3}{k_2+k_3} & = & \alpha_0 \\ \frac{K_1 k_2}{k_2+k_3} & = & \alpha_1 \\ k_2 + k_3 & = & \beta_1 \end{cases} \tag{2.11}$$

If in the model we include a rate constant for the transfer of the tracer from C_2 to C_1, k_4, the spectrum changes its configuration and what was the trapping component becomes a new equilibrating component (Fig.

2.1.2D). This behaviour is consistent with the change of the model configuration shown in Figure 2.1.2C. Obviously the relationship between the αs and βs of the new spectrum with the rate constants is different from the one expressed in Eq. 2.11. The new equivalence can be defined from the solution of the new model.

2.1.3 SA AND TISSUE KINETIC HETEROGENEITY

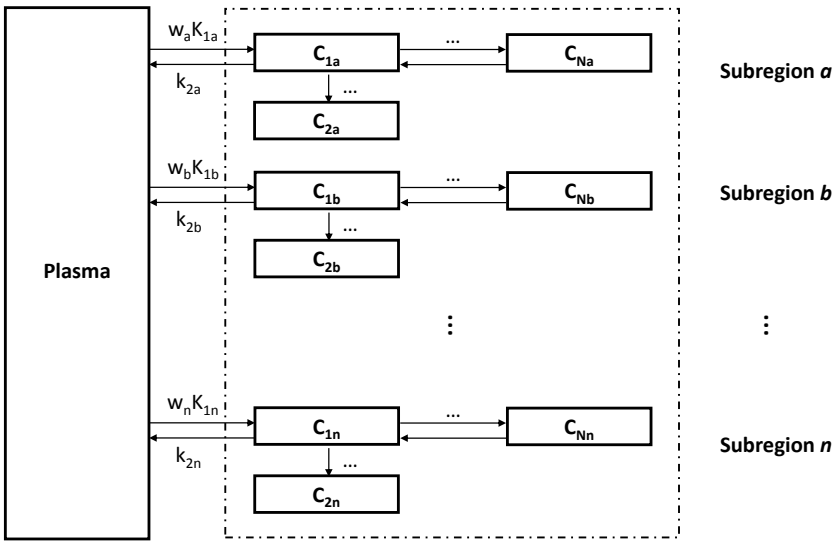

Figure 2.1.3: An example of heterogeneous system. The figure shows the general kinetic model for a heterogeneous system composed by n kinetically homogeneous subregions. Each subregion is then composed by N compartments connected with each other; the configuration can change between subregions. The variables $w_a, w_b, \cdots, w_n$ indicate the weight of each subregion within the observed volume of interest, by assuming that $w_a + w_b + \cdots + w_n = 1$.

Because SA does not require the number of compartments to be fixed *a priori,* the SA technique applies to heterogeneous as well as the homogeneous tissues without any additional assumptions. This characteristic is particularly useful for tracer kinetics studies where the limited spatial

resolution of the PET scanner produces a heterogeneous mixture of kinetically dissimilar tissues within the field of view. In PET brain studies, for example, it not uncommon to have a combined signal from gray and white matter, especially in cortical regions. However, due to the additivity of the spectral components in Eq. 2.1, this effect can be taken into account. In fact, whenever a homogeneous tissue model can be appropriately analyzed with SA, its heterogeneous counterpart can be modelled as well (Schmidt and Turkheimer, 2002).
In the case of heterogeneity, however, the parameters of interest K_1, K_i and V_T resulting from the estimated spectrum assume a different meaning. Considering the general case of a system with n kinetically homogeneous subregions, the activity of the tissue as a whole can be expressed as a convex combination of activity in its subregions $(a, b, \ldots, n)$ (Fig. 2.1.3). It follows that the total activity in the observed volume is given by

$$C_{measured}(t) = (1 - V_b) \cdot \{w_a C_{T,a} + w_b C_{T,b} + \cdots + w_n C_{T,n}\} + V_b \cdot C_b(t) \quad (2.12)$$

where $C_{T,i}$ represents the tracer activity in the i-th subregion.
From Eq. 2.12 it can be easily derived that the kinetic spectrum of the heterogeneous system corresponds to the sum of the kinetic spectra of its subregions. In this case, independently from the number of compartments and the reversibility/irreversibility of each subregion, the sum of estimated components ($\sum_{j=0}^{M} \alpha_j$) as well as the trapping one (α_0) represent the weighted averaged influx rate constant and weighted average trapping of the heterogeneous system. A different treatment has to be reserved for the distribution volume. In fact, only in the case that each subregion presents reversible kinetics, the result of Eq. 2.5 indicates the weighted average distribution volume. On the contrary, in presence of at least one irreversible subregion, no more can be said without the complete modelling of the system. This last condition is however very uncommon among the PET tracers, since the reversible/irreversible kinetic behaviour of a specific tracer tends to be maintained between different biological tissues.

A great advantage of applying SA to inhomogeneous systems is the possibility of detecting the degree of heterogeneity. From SA analysis it is possible to compute the heterogeneity fraction (HF), defined as the number of homogeneous distinguishable subregions within the investigated system. As consequence of the indistinguibility theorem, however, it is important to note that the computation of HF can be performed only when the system is completely characterized and the corresponding compartmental model fully-described.
Since the problem of heterogeneity involves not only the modelling approach used to quantify the data but also the structural limits of imaging instruments as well as the measurement noise, it will be more exhaustively discussed in Chapter 5.

2.1.4 SA AND MODEL DEVELOPMENT

Alternative to tissue kinetic quantification, SA has been used as tool for the compartmental model development of several PET tracers. Examples of this can be found in (Bertoldo et al., 1998, 2001, 2005; Myers et al., 2012). This approach is particularly useful when new PET tracers are analyzed for the first time. In this context, SA offers the possibility of deter-

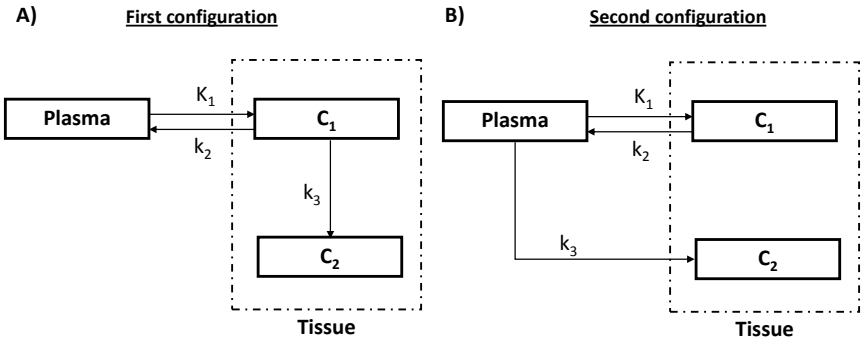

Figure 2.1.4: Two compartmental models with equivalent kinetic spectrum. Although they have different configurations both methods are compatible with the same spectrum defined by a trapping and an equilibrating component.

mining the number of compartments present in a system as well as their

types, distinguishing between equilibrating components and trapping. However, as discussed in the previous paragraph, it cannot say anything about their interconnections, and without any information about the system-of-interest physiology it is impossible to determine an unequivocal correspondence between the spectrum and its equivalent model. Thus, for a particular estimated spectrum it is only possible to associate a class of equivalent compartmental representations which have in common the same number of compartments (Fig. 2.1.4). Then, using the physiological knowledge of the system, it is possible to choose from among the class alternatives the optimal compartmental configuration for the description of the kinetics of the tracer under study. This procedure is theoretically always applicable but may not be advisable for real practice. It can happen very often that the presence of noise in the data (especially for high level of error) produces a biased number of SA estimated components (generally higher than the true value) leading toward an erroneous class of model configurations.
For this reason when the purpose of SA is to find the number of exponential components of the system impulse response without having to define a structure beforehand, defining a grid of eigenvalues β_j and estimating only the relative amplitudes α_j is not necessary. The right approach to follow is simply to estimate M the number of exponentials necessary for the SA model (Eq. 2.1) to give a good fit to the data by using models of increasing order. For instance, one can start first with a one-exponential model

$$C_T(t) = \alpha_1 \cdot C_p(t) \otimes e^{-\beta_1 t} \tag{2.13}$$

and estimate by Non Linear Least Squares (NLLS) the parameters. Then trying a two-exponential order, with one trapping component

$$C_T(t) = \alpha_1 \cdot \int C_p(t) + \alpha_2 \cdot C_p(t) \otimes e^{-\beta_2 t} \tag{2.14}$$

and then with two equilibrating components

$$C_T(t) = \alpha_1 \cdot C_p(t) \otimes e^{-\beta_1 t} + \alpha_2 \cdot C_p(t) \otimes e^{-\beta_2 t} \tag{2.15}$$

and so on. The standard model parsimony criteria techniques (Hirotugu, 1971) can be used to choose the best model. This approach firstly presented by Bertoldo and colleagues (Bertoldo et al., 1998) is also called nonlinear spectral analysis (NLSA). Compared to standard SA it offers several advantages. First of all not only the standard deviation error of the αs estimates, but also the precision of the βs are provided. In fact, if βs are selected from a predetermined grid, as in the standard SA approach, it is impossible to obtain a measure of their precision. This information can be combined with the parsimonia criteria for the selection of the best model. Secondly, estimation of the β_j within a pre-fixed compartmental structure avoids the problem of the extra components as in the standard SA. For all these reasons NLSA represent the most appropriate SA approach for model identification.

2.2 Linear and filtered solutions

2.2.1 Exponential Spectral Analysis (ESA)

The first applications of the SA model (Cunningham and Jones, 1993), were done with brain PET datasets, specifically for the evaluation of cerebral blood flow, cerebral glucose utilization and for opiate receptor ligand binding. $H_2{}^{15}O$, [^{18}F]FDG , and [^{11}C]DPN PET data were considered for this purpose. The definition of the βs grid was simply based on a logarithmical distribution, while the estimation was solved with the simplex method (Cunningham and Jones, 1993).
After these attempts, the SA model has been widely used in a large variety of testing conditions, although with different solutions for the β grid and for the implementation of the nonnegative estimator. Despite these dissimilarities, it is very common to refer to this class of methods in the same way, as Exponential Spectral Analysis (ESA). ESA has been applied to preclinical (rats and rabbits) (Bentourkia, 2003; Marshall et al., 2004) as well as clinical data. Most of its applications are related to the investigation of brain tissues (Edison et al., 2009; Grunder et al., 2001; Hammers et al.,

2005, 2007; Miederer et al., 2009; Myers et al., 2012; Ramlackhansingh et al., 2011; Richardson et al., 1996; Turkheimer et al., 1994, 1998, 2007; Weeks et al., 1997; Wester et al., 2000; Willoch et al., 2004) even if ESA has been applied to PET studies involving heart (Bertoldo et al., 1998), skeletal leg muscle (Bertoldo et al., 2001; Pencek et al., 2006), breast cancer (Kenny et al., 2005; Tomasi et al., 2011b, 2012) and gastrointestinal cancer (Wells et al., 2003). Most of these applications aimed to provide reliable and robust alternatives to the standard quantification methodologies.

2.2.2 Numerical filtering

The ESA method is well known to be sensitive to noise in the data, with the bias being highly dependent on the level of noise present (Cunningham et al., 1998; Gunn et al., 2002; Turkheimer et al., 1994, 1998). Over the years several strategies to lessen the impact of noise in the data on estimated spectra and on the parameters of interest have been proposed in literature.
The first attempt to solve the problem was done by including penalty functions on the NNLS algorithm (Cunningham et al., 1993). However, the definition of an adequate penalty function to use in practice has proved to be difficult. This concept was further developed by Gunn and colleagues in 2002 and it originated what are now called basis pursuit methods (Gunn et al., 2002; Peng et al., 2008).
In 1994, Turkheimer and colleagues proposed an high-pass filter for equilibrating components, with the aim of improving estimates of α_0 and thus determining a more accurate and precise regional cerebral metabolic rate for glucose in [^{18}F]FDG PET studies (Turkheimer et al., 1994). The filter assumes that equilibrating components detected with $0 < \beta_j < \beta_{cutoff}$ result from noise in the data, shifting trapping components from their true position at $\beta_j = 0$. The filter was effected by subtracting from the measured ROI TAC all components with $\beta_j \geq \beta_{cutoff}$; the remainder, which was assumed to equal the integrated component $\alpha_0 \int C_p(\tau)d\tau$, was used to re-estimate α_0. Even though the use of this filter have been shown to produce better estimates of α_0 than the standard ESA, it rep-

resents an incomplete solution. In particular this method does not account for the phantom components at intermediate and high frequency (Veronese et al., 2010), which, as well the low-frequency components, contribute to the description of tissue tracer kinetics. This may lead to biased estimate for K_1 and V_b (Veronese et al., 2010), indicating that a correct SA-based quantification requires a correct estimation of the whole spectrum. In 1993 and 1997, Takodoro and Weeks (Takodoro et al., 1993; Weeks et al., 1997) proposed to use the SA impulsive response function, $IRF(t)$, as an alternative to the SA spectrum to compare studies. In particular they suggested to use the $IRF(t)$ value at early times of the PET study for the analysis of tracer delivery and at the late times for tracer retention. However, even if $IRF(t)$ parametric imaging proved to be more robust to the noise, a quantitative formalization about the relationship between this approach and the underlying kinetic processes is missing.
In 1998 two different alternatives were proposed. First of all a bootstrap approach based on residual resampling was implemented to simulate the effect of the noise on the SA spectrum and to correct for possible bias (Turkheimer et al., 1998). The bias-corrected bootstrap spectrum was shown to be a good estimator of the average spectrum that would have been obtained if repeated samples of the measured data were available. In addition the technique was shown to require minimal assumptions about noise in the measurements and thus to have general applicability. However, because the procedure was computationally intensive, its applicability is limited in the case of very large datasets such as in voxel-wise PET analysis. In the same year, a new procedure for the suppression of noise artefacts on the estimated SA spectrum was proposed by Cunningham and colleagues (Cunningham et al., 1998): standard statistical tests and information criteria were applied to subselect the true estimated components from a given SA spectrum, removing those related to the noise in the data. In the same work, some important considerations about the spacing of the basis functions to be used in SA were presented (Cunningham et al., 1998).
Parallel to the developing of new SA improved versions, some efforts

were done to denoise the PET images, exploiting the idea that quantification can be improved by increasing the SNR of the analyzed data. This preprocessing step to smooth the data then allows subsequent analysis by methods such as SA to be substantially improved in terms of their mean squared error. Several solutions have been proposed in literature, but specifically to SA methodology, the most significant attempts were represented by the use of *wavelets* (Turkheimer et al., 1999, 2000) and by the *functional smoothing approach* (Jiang et al., 2009). Despite of their great potential, both of these approaches have not significantly impacted the PET modelling community.

At the present time, the most successful filtered version of the SA method is represented by Rank-shaping spectral analysis (RS) (Turkheimer et al., 2003). RS is a Bayesian development of standard SA, optimized for reversible tracers, which works on the same principles as SA but without requiring the non-negativity constraint in the estimator. In fact, to overcome the high sensitivity of the standard SA to the noise in the data, RS implements a Kalman filter of the estimated kinetic spectrum providing reliable estimates of V_T (for the derivation of this method, the interested reader is referred to the original reference by Turkheimer and colleagues (Turkheimer et al., 2003)). Unlike other SA-based methods, RS can be applied with an arterial input function as well as with a reference region. RS has been shown to be a precise and accurate quantification method (Hammers et al., 2007; Turkheimer et al., 2003, 2007, 2012), demonstrating to be a reliable tool for parametric imaging, even in high-noise conditions (Turkheimer et al., 2003). On the other hand, RS applicability is limited to the reversible tracers only. This is an important restriction for PET quantification because there are several PET tracers with irreversible trapping. The most relevant example is represented by [^{18}F]FDG , one of the most used PET tracers in research and clinical practices[4]. It is mainly for this reason that a complementary version of RS for irreversible tracers

[4]There are more than 15-thousand works registered in Pubmed (`http://www.ncbi.nlm.nih.gov/pubmed`) database involving [^{18}F]FDG PET tracer. Among these works almost 3-thousand refer to 2012.

needs to be developed.
In the next section we will introduce SAIF (*Spectral Analysis with Iterative Filter*) (Veronese et al., 2010), a filtered SA-based algorithm for the quantification of PET tracers with irreversible uptake.

2.3 A NOVEL SA METHOD FOR IRREVERSIBLE TRACER: SAIF

2.3.1 THE ALGORITHM

In irreversible tracers, the quantification of the trapping and the tracer transport (from plasma to tissue) requires good estimates of both α_0 and $\sum_{j=0}^{M} \alpha_j$. From these parameters it is then possible to derive tracer-specific parameters of interest like the regional glucose utilization, in case of [^{18}F]FDG PET studies (Turkheimer et al., 1994), or the rate of cerebral protein synthesis, as in case of L-[1-^{11}C]Leucine brain PET studies (Veronese et al., 2010). To do this, it is necessary to separate low frequency equilibrating components from trapping as well as to separate high frequency equilibrating components from blood. If the latter is not done successfully, K_1 will be substantially overestimated.
The starting point for definition of a new filter was the high-pass filter of Turkheimer (1994) (Turkheimer et al., 1994). Instead of considering the components with $0 < \beta_j < \beta_{cutoff}$ as resulting from trapping components and noise in the data, the new filter allows that these components may also be due, at least in part, to slowly equilibrating components present in the system. Therefore, the effect on remaining equilibrating components of removing components with $0 < \beta_j < \beta_{cutoff}$ must be considered. The second aspect of the new filter was setting an upper cut-off frequency above which equilibrating components are assumed to be indistinguishable from blood in brain. Removal of these components from the spectrum affects estimates of remaining equilibrating components as well as blood volume. The filtering includes two steps that are repeated until stabilization of Weighted Residual Sum of Squares (WRSS) is reached. Because of the iterative cycle, this new SA algorithm is called *Spectral*

Analysis with Iterative Filter" (SAIF). The procedure is as follows:

1. ***Define the grid*** *for β_j, $j = 1, 2, ..., M$ and perform spectral analysis without filtering.*

2. ***Set up Bandpass filter****: Define an equilibrating component passband [β_L,β_U] such that all components with $\beta \in$[β_L,β_U] are considered to be real equilibrating components of the system.*

3. ***New estimates of trapping and blood volume****: Assume all components with $0 < \beta_j < \beta_L$ result from noise having shifted some or all of the trapping components from their true position in the spectrum, and all components with $\beta_j > \beta_U$ are shifted blood components. Subtract components inside the passband from measured activity ($C_{measured}(t)$) and estimate new values of the coefficient of the trapping compartment (α_0) and the blood volume fraction (V_b).*

4. ***New estimates of equilibrating components****: Recognizing that components outside the passband may have at least partially resulted from real equilibrating components that were shifted due to noise, re-estimate equilibrating components as follows. Using the new values of α_0 and V_b from Step 3, subtract the contribution of the trapping and blood components from the measured data, and re-estimate the set of equilibrating components using NNLS with the grid restricted to $\beta \in$[β_L,β_U].*

5. ***Stop Criterion:*** *Calculate the difference between WRSS before after Step 3 and after Step 4; if it is less than $\varepsilon = 0.0001$ stop . Otherwise restart at Step 3.*

Using change in WRSS as the stop criterion allows the data to auto-determine the number of iterations. When all equilibrating components initially identified with spectral analysis (before filtering) are inside the passband, only one iteration is required. A schematic summary of the SAIF procedure is reported in Figure 2.3.1.

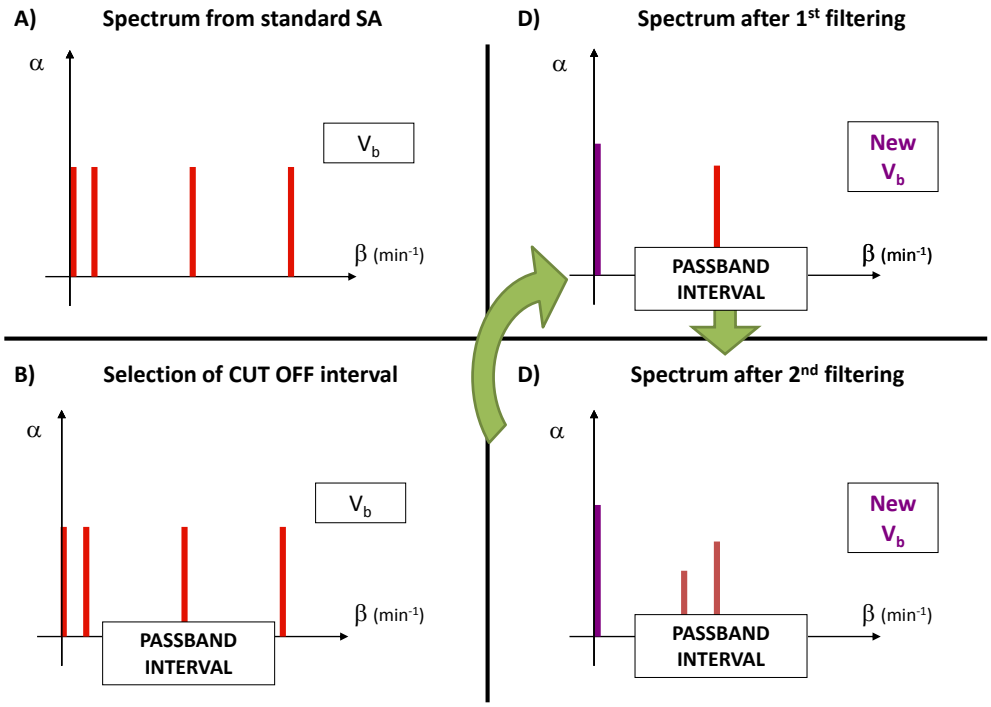

Figure 2.3.1: SAIF filtering procedure. (A) The spectrum provided by ESA; (B) The choice of the cut-off interval on the evaluated spectrum; (C) 1st filtering: The contributions of components inside the cut-off interval are subtracted from the data and new values of the trapping and V_b are estimated; (D) 2nd filtering: The contributions of the trapping and blood volume are subtracted from the data and new values for the components inside the cut-off interval are estimated. The iteration cycle is the repetition of the last two points. The WRSS variation provides the stopping criterion.

2.3.2 SAIF IMPLEMENTATION

As all the SA methods, implementation of SAIF requires 1) the definition the βs grid, 2) the selection of weighting factors for nonnegative least squares estimation, 3) the correction for the artefacts that arise from the discrete, non-continuous nature of the grid, and 4) the computation of the standard errors in the parameter estimates. In addition to these elements, SAIF needs for the passband filter interval, i.e. [β_L,β_U].
Since the implementation choices can impact the performance of the method, we decided to follow the literature suggestions for its realization. The

logarithmic distribution of β_j, $j = 1 : M$ was chosen according to (DiStefano, 1981; Turkheimer et al., 1994). The lower limit of the distribution was defined as $\beta_1 = 1/(3T_{end})$ where T_{end} is the end time of PET study, and the upper limit was provided by $\beta_M = 3/T_{in}$ where T_{in} is the duration of the first PET scan of the study (Turkheimer et al., 1994). One hundred points were included ($M = 100$). Therefore:

$$\beta_j = \frac{1}{\tau_j} \tag{2.16}$$

and

$$\tau_j = \tau_{j-1} \cdot \left(\frac{T_{end}}{T_{in}}\right)^{\frac{1}{1-M}} \tag{2.17}$$

with $j = 2 : \cdots : M$. The $M + 1$ unknown values α_j were estimated with the nonnegative linear least squares algorithm implemented in Matlab® (The MathWorks, Inc, Natik, MA, USA) by the *lsqnonneg.m* function. Weights were assumed proportional to the variance of the decay-corrected measured activity in the region, which was differently defined depending on the characteristics of the tracer. In all the cases a relative weighting scheme was applied, with a scale variance factor (σ^2) estimated *a posteriori* from the WRSS (Carson et al., 1983). A lower bound on the standard deviation of the estimated α_j was determined from the inverse of the Fisher Information matrix, following the Cramer-Rao Theorem (Cramer and Herald, 1946).

Due to the discrete nature of the grid of βs, the algorithm cannot always place a component in the exact position that would yield the best fit of the data; it is instead constrained to use only βs included in the grid. When an adequate approximation of the optimal β is not on the grid, the algorithm chooses instead two consecutive values of βs that bracket the optimal value. Under the assumption that consecutive components resulted from this phenomenon rather than from two nearly equal components in the data, we replaced consecutive components pairs $Comp_j$ (β_j;α_j) and $Comp_{(j+1)}$ ($\beta_{(j+1)}$;$\alpha_{(j+1)}$) with a new, single component $Comp_{j*}$(β_{j*};α_{j*})

calculated from the amplitude-weighted average as follows:

$$\alpha_{j*} = \alpha_j + \alpha_{j+1} \tag{2.18}$$

$$\beta_{j*} = \frac{\beta_j \cdot \alpha_j + \beta_{j+1} \cdot \alpha_{j+1}}{\alpha_j + \alpha_{j+1}} \tag{2.19}$$

2.3.3 APPLICABILITY POTENTIAL

For the characteristics of the SA model and the way in which it has been implemented, SAIF presents general applicability to all the PET tracers with irreversible uptake. Coherently with all the SA-based methods, also SAIF is adequate only for the description of single-input/single-output noncyclic biological systems. In terms of parameters of interest, SAIF returns the information K_i and K_1 rate constants (as expressed by Eq. 2.3 and 2.4), as well as the blood volume (V_b). In addition, it provides the whole spectrum of the kinetic components indicating the number of compartments necessary to describe the data.
To show the potential offered by the method, the next chapters present some representative examples of SAIF application to several human PET studies performed using different tracers. The analyzed datasets were specifically chosen with the aim of testing SAIF under different conditions. For a clearer presentation, the results related to brain and non brain studies will be separately reported.

Never trust anything that can think for itself
if you can't see where it keeps its brain.

J. K. Rowling

3

Applications to brain studies

THIS CHAPTER PRESENTS some applications of SAIF in simulated and measured brain PET data. For the characteristic of irreversibility required by the method, SAIF was tested in L-[1-^{11}C]Leucine and [^{11}C]SCH442416 PET datasets its performance compared with those offered by the current state of the art presented in literature.

3.1 BRAIN SIMULATED DATASETS

3.1.1 DEFINITION OF SIMULATED DATA

Previous to measured data analysis, SAIF was tested on a representative set of simulated datasets representing both regional and voxel-wise noise level. The following subsection summarize the main details of simulation procedure. Firstly, noise-free tissue kinetic TACs ($C_{tissue}(t)$) were gener-

ated by the convolution between a pre-defined measured arterial input function (C_p) with a selection of simulated kinetic spectra as in Eq. 2.1. C_p was computed as the population average of 50 manually-measured arterial input functions derived from 50 L-[1-^{11}C]Leucine PET studies. For the selection of kinetic spectra, instead, we chose a variety of results obtained from the analysis of several brain PET studies performed with irreversible tracers (Bertoldo et al., 1998; Bishu et al., 2008). A total number of 75 spectra were generated. The correspondent parameters of interest (K_1, K_i and V_b) derived from the simulated spectra are reported in Table 3.1.1. Noise-free simulated TACs $C_{sim}(t)$ were generated according PET

Table 3.1.1: Simulated parameter values

	K_i	K_1	V_b
Index	(ml/cm^3/min)	(ml/cm^3/min)	(unitless)
Mean	0.026	0.078	0.071
SD	0.031	0.060	0.015
min	0.006	0.025	0.040
max	0.100	0.200	0.104

measurement equation (Chapter 2, Eq. 2.7), here recall

$$C_{sim} = (1 - V_b) \cdot C_{tissue}(t) + V_b \cdot C_b(t) \tag{3.1}$$

where $C_b(t)$ represents the concentration of the tracer in the whole blood including radiometabolites. For consistency with arterial data, $C_b(t)$ curve was derived through the same procedure used for the definition of $C_p(t)$. A total number of 75 noisy-free simulated TACs, one per simulated kinetic spectrum, were computed (Fig. 3.1.1A). Noisy-TACs data were then generated by the sum of noise-free data and simulated white noise (zero-mean and Gaussian distribution). Data were sampled according to the PET acquisition protocol defined in (Bishu et al., 2008), i.e. 42 frames of data (16x15, 4x30, 4x60, 4x150, 14x300 seconds). The simulated noise variance was defined as it is most frequently done in PET, i.e. according to

the formula proposed by Mazoyer and colleagues (Mazoyer et al., 1986):

$$Var(t) = \sigma^2 \cdot \frac{(C_{sim}(t))}{\Delta t} \tag{3.2}$$

where σ^2 represents the scale factor for the error variance, $C_{sim}(t)$ represents the simulated noise-free TAC and Δt represents the lengths of scan intervals. In order to test the performance of SA methods at both region and voxel levels, high and low SNR scenarios were generated by modulating the simulated noise for different values of σ^2. More than 2 orders of magnitude were used to simulate voxel SNR level respect to the regional one. For each simulated kinetic, 100 noisy realizations were generated for both high and low SNR scenarios, for a total of $15,000$ simulated noisy TACs. A representative set of the signals obtained for each scenarios is reported in Figure 3.1.1 (B and C).

3.1.2 Method implementation and performance indexes

For comparative purposes, simulated data analysis was performed with both ESA and SAIF algorithms. The same grid of βs and the same weighting procedure, as defined in Section 2.3.2 were applied for consistency. SAIF passband interval bounds were respectively defined as the minimum and the maximum of the simulated equilibrating β components ([β_L,β_U]=[0.02 min^{-1}, 0.3 min^{-1}]). The same passband was maintained for the analysis of both low and high SNR data. Trapping rate constant (K_i), average influx rate constant (K_1) and blood volume (V_b) were considered as parameters of interest. For each noise level and each parameter of interest p, the mean percentage bias ($\%bias(p)$) was computed as:

$$\%bias(p) = \sum_{s=1}^{75} \left[\frac{1}{75} \sum_{k=1}^{100} \frac{1}{100} \left(\frac{p_{k,s} - \hat{p_{k,s}}}{\hat{p_{k,s}}} \right) \right] \tag{3.3}$$

where $p_{k,s}$ represents the k-th estimated values of the s-th simulated kinetic spectrum and $\hat{p_{k,s}}$ represents the k-th simulated reference values of

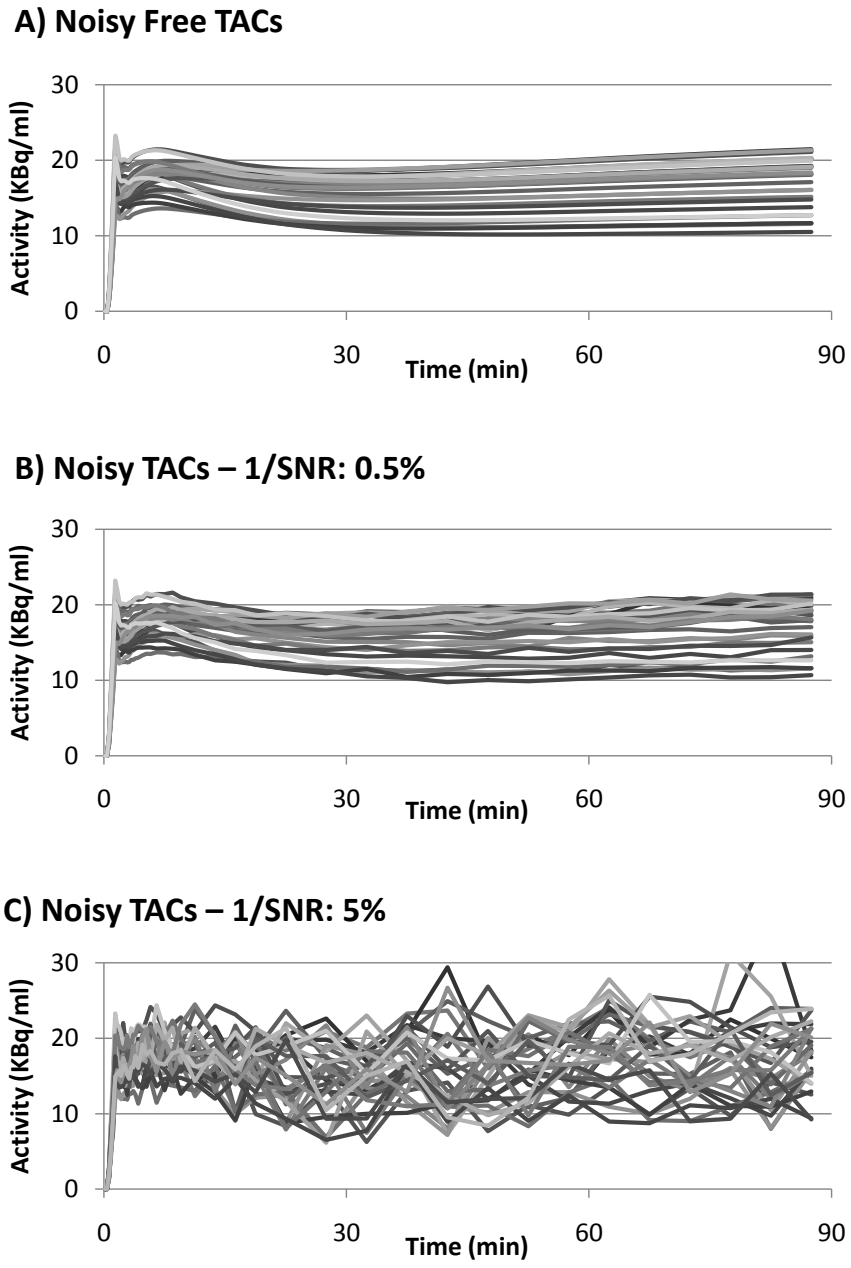

Figure 3.1.1: Simulated dynamic PET TACs. Panel A shows noise-free C_{tissue} TACs. Panel B and C show some representative noisy realizations of dynamic PET data at high and low SNR level.

the s-th simulated kinetic spectrum. To note that k ranges from 1 to 100 (number of simulated noisy TACs per spectrum) and s ranges from 1 to 75 (total number of simulated spectra). To complete the method performance analysis, ESA and SAIF accuracies were evaluated through the calculation of $\%bias$ SD and *min-max* range.
In addition to $\%bias$ characterization, for each tested method, the successful rate (SR) was computed as the fraction of the successfully estimates over the total number of simulations[1].

3.1.3 RESULTS

Table 3.1.2: Simulated parameter values

	ESA					**SAIF**				
Parameters	Bias K_i (%)	Bias K_1 (%)	Bias V_b (%)	WRSS (unitless)	SR (%)	Bias K_i (%)	Bias K_1 (%)	Bias V_b (%)	WRSS (unitless)	SR (%)
				Region Level						
Mean	-4	-4	-4	1.6	100	1	1	1	2.2	100
SD	1.9	1.9	1.9	2.7	0	2.5	2.5	2.5	4	0
Min	-6.3	-6.3	-6.3	0.3	100	-0.6	-0.6	-0.6	0.3	100
Max	-0.2	-0.2	-0.2	8.7	100	9.1	9.1	9.1	13.8	100
				Voxel Level						
Mean	-13	-13	-13	236	86	0	0	0	239	94
SD	6	6	6	197	7	6	6	6	199	6
Min	-22	-22	-22	88	73	-7	-7	-7	89	84
Max	-2	-2	-2	723	95	17	17	17	736	100

Table 3.1.2 summarizes the main results obtained by ESA and SAIF in the simulations. In particular it can be observed that:

- the increase of noise level produced a worsening of precision an accuracy in both methods for all the parameters of interest. This behaviour was quite expected, and it is consistent with one of most common engineering principles for which the higher the noise, the more difficult the measure. Also the number of outliers results influenced by noise level as indicated by the decrease of SR from ROI to voxel analysis;

[1]We considered *successfully estimates* all the cases in which the algorithm converged to a physiologically-acceptable solution.

- compared to ESA, SAIF provided higher precision and accuracy for most of the parameters of interest. In addition SAIF mean SR was found always equal or lower than the ESA one, indicating a superior robustness of the method to the testing conditions;

- compared to SAIF, ESA provided lower WRSS. This is consequence of the filtering procedure implemented in SAIF.

Since the results in Table 3.1.2 might be mitigate by the average effect, we considered also the whole distribution of $\%bias$, i.e. the distribution of $\%bias$ obtained comparing each simulated and estimated TACs. Figure 3.1.2 reports the $\%bias(K_i)$ distributions obtained with ESA and SAIF (selecting the 90% of confidence interval). Both low and high noise scenarios are visualized. ESA results were comparable with those already presented (Turkheimer et al., 2003; Veronese et al., 2010), which had highlighted the high sensitivity of the ESA method to the noise in the data: at high SNR level $\%bias(K_i)$ was limited in both term of mean and variability ($\%bias(K_i)$ mean±SD: -1.4%±4.9%) while at low SNR level it is evident a worsening of its performances ($\%bias(K_i)$ mean±SD: -4.3%±20.0%). On the other hand, SAIF demonstrated limited mean $\%bias(K_i)$ (for high SNR 0.35%; for low SNR 1.5%) and variability consistent with the level of noise of the analyzed data (SD $\%bias(K_i)$ for high SNR 1.9%; SD $\%bias(K_i)$ for low SNR 9.6%). Similar results were obtained also for the other parameters of interest.

A visual but very intuitive comparison of ESA and SAIF performances at high noise level can be done by analyzing the results reported in Figure 3.1.3. The 75 simulated spectra were associated to 75 different regions of a MNI brain atlas, derived from (Hammers et al., 2003). The correspondence between spectra and regions was done only for graphical purpose, without any precise relationship between simulated kinetic and the physiology of the anatomical region. In this way, however, it was possible to identify each ROI with a different vector of simulated parameters (K_i, K_1 and V_b) (Fig. 3.1.3A). Then, like a MonteCarlo simulation, all the vox-

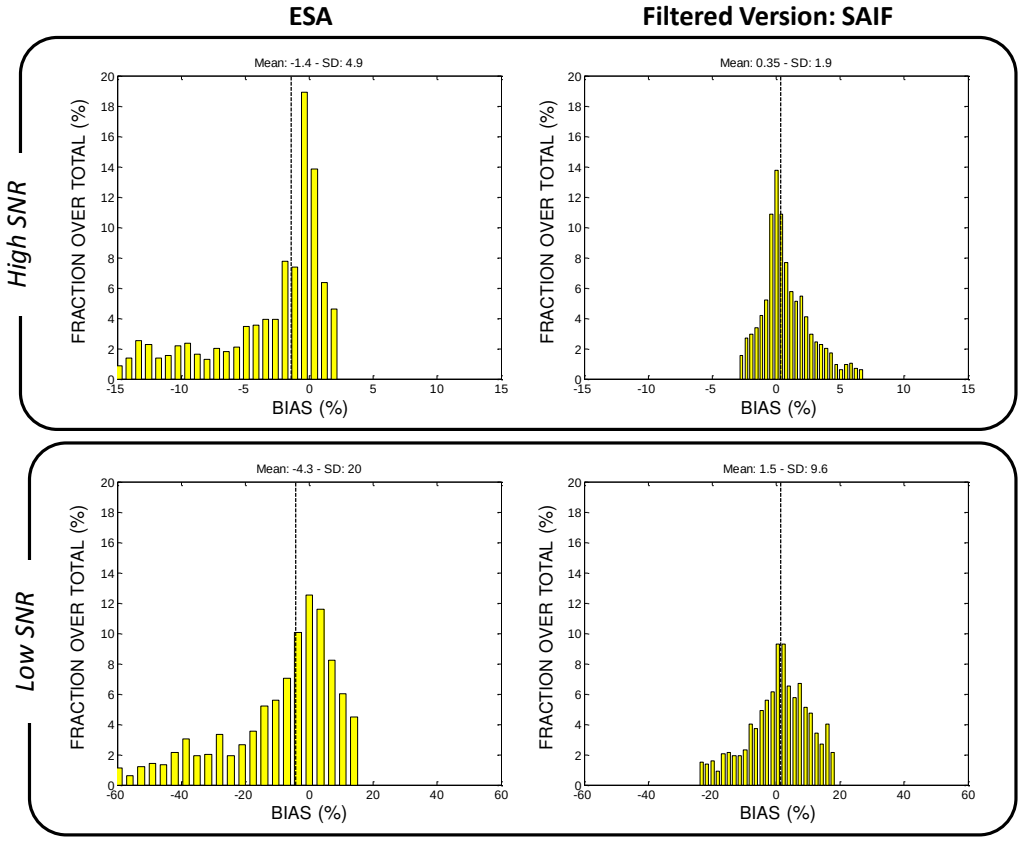

Figure 3.1.2: $\%bias(K_i)$ distribution obtained with ESA and SAIF in simulation. The figure shows the single-estimate K_i bias distribution at high and low SNR levels (top and bottom of the panel respectively).

els within a given ROI were associated with different noise realizations of the same simulated kinetic activity correspondent to specific region. The results obtained by applying ESA and SAIF to this dataset are reported in Figure 3.1.3B,C. It is clearly evident how ESA K_i results are completely out of the range of the simulated values. On the contrary, SAIF presents K_i estimates in agreement with the reference values. The fraction of SAIF failures and outliers was $<1\%$ respect the total number of simulated TACs. The presence of *salt and pepper* disturb on the SAIF K_i maps (Fig. 3.1.3C), consequence of the effect that the simulated noise had on SAIF quantification, represents a good indication of unbiased behaviour of SAIF.
Figure 3.1.4 shows the SAIF $\%bias(K_i)$ as function of simulated kinetics (randomly chosen from the previous simulated spectra, Table 3.1.2) and a set of arterial input curves (AIF) (Fig. 3.1.4A). For this purpose, 50 AIFs were extracted from several multicentre PET studies, referring to manual and automatically sampling. This choice was made to be as much general and tracer independent as possible. Analysis of SAIF performance in relationship to AIF was performed at both high and low SNR levels.
In agreement with previous results, it can be noticed how the increase of noise is coupled with a decrease of estimate precision and accuracy (Fig. 3.1.4B,C). The same figure shows also how the SAIF performances are quite independent from the characteristics of AIF used, while they can be correlated with the choice of spectral kinetic: it appears evident from the matrix analysis (Fig. 3.1.4B,C) that some rows result more biased than others, indicating a higher sensitivity of the method to the particular simulated spectrum. Since a unique SAIF passband was set, it might result that this choice was not adequate for all the simulated kinetic spectra.

3.1.4 CONSIDERATIONS

The ESA simulation results were found in agreement with those presented already in literature, which had highlighted the necessity of fil-

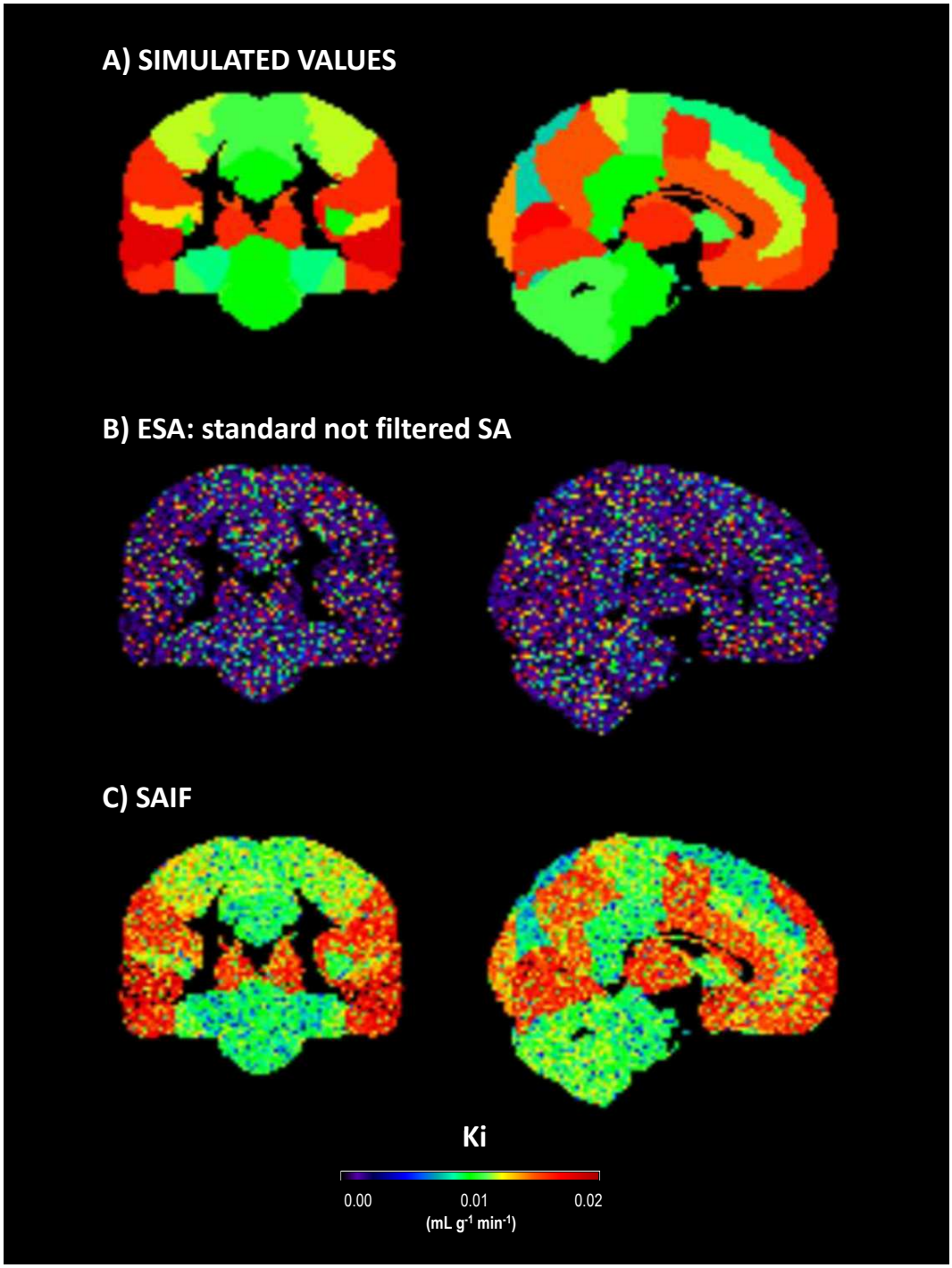

Figure 3.1.3: Parametric maps of $\%bias(K_i)$ obtained with ESA and SAIF in simulation. Panel A-C report respectively simulated, ESA-estimated and SAIF-estimated K_i parametric maps, obtained from high-noise simulated PET data.

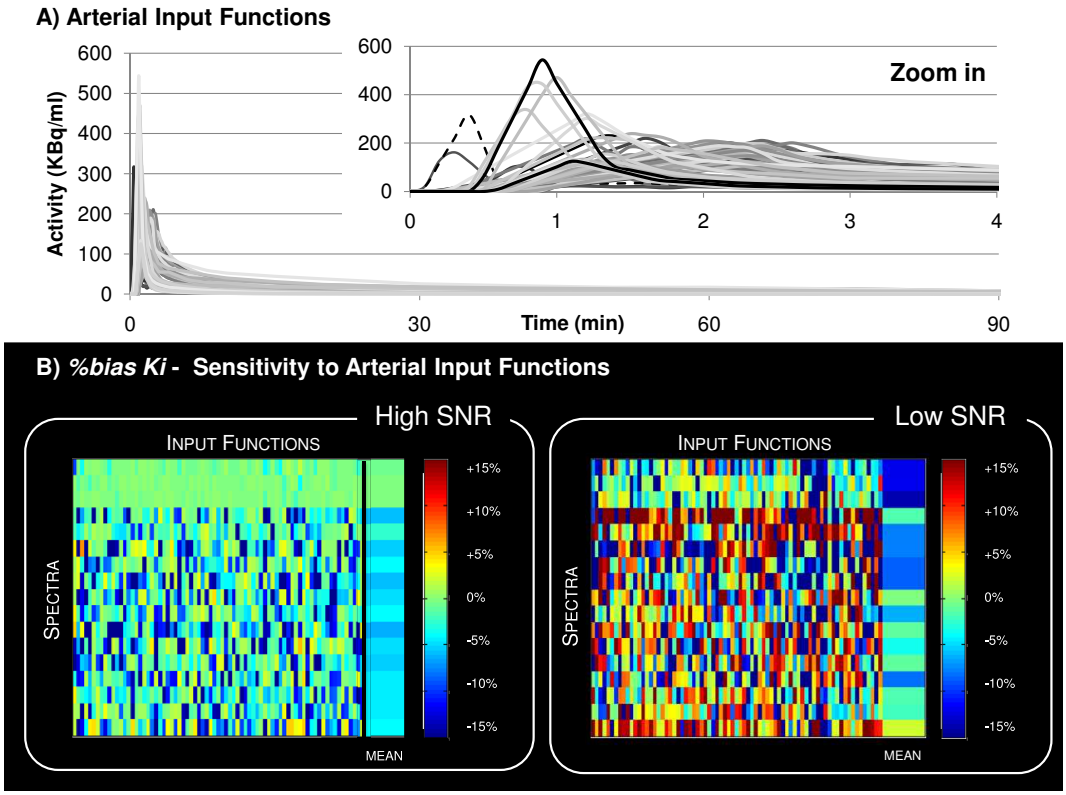

Figure 3.1.4: Sensitivity of SAIF performance to the Arterial Input Function. Panel A shows the time-activity course of the AIFs used for the simulation. Panel B and C report the $\%bias(K_i)$ obtained for different simulated kinetic spectra and different AIFs. Both regional and voxel noise level were simulated.

tered version of SA to improve the robustness of the method to the presence of noise (Section 2.2.2). On the contrary, SAIF proved to be a robust and reliable quantification method, providing precise and accurate estimates with limited number of outliers even at high noise conditions. Although a methodological validation cannot be limited to simulated data analysis, these results represent a good starting point for SAIF development and its application in real measured PET data.
To note that SAIF performances are heavily dependent from the choice of the passband filter. In this simulation it was possible to define the SAIF filter bounds accordingly with the characteristics of the simulated kinetic spectra. However, when SAIF has to be applied to measured data, this information is not available and other strategies need to be found for the passband filter definition.

3.2 L-[1-^{11}C]Leucine PET studies in healthy subjects

In this section the main results related to SAIF application to L-[1-^{11}C]Leucine PET images are presented. Historical, SAIF was developed for region-wise quantification of this data (Veronese et al., 2010). the, since it showed to be robust to the presence of noise in the data, its use was extended for the voxe-lwise analysis (Veronese et al., 2012b).

3.2.1 Introduction

The L-[1-^{11}C]Leucine PET method was developed and validated in order to make possible fully quantitative measurement of regional rates of cerebral protein synthesis ($rCPS$) in human subjects (Schmidt et al., 2005; Smith et al., 2005). Biosynthesis of proteins is a fundamental process necessary for physiological maintenance and functioning of all organisms and its alteration is related to a large set of clinical disorders (Collins et al., 1980; Qin et al., 2005; Smith and Kang, 2000; Widmann et al., 1991, 1992). $rCPS$ also changes during brain development (Sun et al., 1995) and normal aging (Ingvar et al., 1985).

In the first fully quantitative L-[1-^{11}C]Leucine PET studies in human subjects (Bishu et al., 2008, 2009; Sundaram et al., 2006) $rCPS$ was determined at the region-of-interest (ROI) level. Because the analysis was based on a homogeneous tissue kinetic model for leucine, heterogeneity within a region was not taken into account. Tomasi and colleagues (Tomasi et al., 2009b) recently developed a method to estimate kinetic model parameters and $rCPS$ on a voxel-by-voxel basis. By averaging the estimated parameters and $rCPS$ over all voxels within a ROI, this method produces estimates for the region as a whole that take into account differences in kinetics among the regions voxels. In simulation studies, estimates derived from the voxel-wise approach were shown to be more accurate than those derived by fitting ROI data to a homogeneous tissue kinetic model (Tomasi et al., 2009b). This effect was mainly caused by the presence variable fraction of kinetic heterogeneity in the data (Bishu et al., 2008, 2009), characteristic that prevented also the application of a fixed compartmental model configuration for ROI level quantification.
For its characteristics of robustness to the noise in the data and applicability to both homogeneous and heterogeneous tissues without any a priori, SAIF might be a valid solution for both regional and voxel-wise quantification of $rCPS$ with L-[1-^{11}C]Leucine PET images.

3.2.2 Dataset and methods of analysis

Data from 12 healthy awake male subjects (age 20-24) from previous study (Bishu et al., 2008) were used. The criteria for subject inclusion and the procedure for L-[1-^{11}C]Leucine PET studies are described in detail in (Bishu et al., 2008). All studies were performed on the ECAT High Resolution Research Tomograph (HRRT) (CPS Innovations, Knoxville, TN), which has spatial resolution of $\sim$2.6 mm full width at half maximum (Wienhard, 2002). The 90-min emission scan was initiated coincident with a 2-min intravenous infusion of 20-30 mCi of L-[1-^{11}C]Leucine (one subject received a 1-min infusion). Estimated leucine specific activity was 3 mCi/nmol. Images were reconstructed using motion-compensated 3D

ordinary Poisson ordered subset expectation maximization (Carson et al., 2003) as 42 frames (16x15, 4x30, 4x60, 4x150, 14x300 sec); voxel size was 1.21x1.21x1.23 mm. Arterial blood sampling was performed; concentrations of unlabeled and labeled leucine in plasma and total ^{11}C and $^{11}CO_2$ activities in whole blood were measured according to methods detailed in (Bishu et al., 2008). All subjects underwent a T1-weighted MRI of the brain for ROI placement. A total of 18 regions and whole brain were evaluated in all subjects except one in which the hypothalamus could not be drawn.

SAIF analysis was performed both at region level (9 subjects) and at voxel level (6 subjects). As parameter of interest, particular attention was given to K_1 (Chapter 2, Eq. 2.3), the fraction of unlabeled leucine in the precursor pool for protein synthesis derived from arterial plasma, λ, ($\lambda = (K_1 - K_i)/K_1$) and $rCPS$ defined as:

$$rCPS = \frac{K_i}{\lambda} \cdot \hat{C}_p \tag{3.4}$$

where $\hat{C}_p$ represents the measured concentration of unlabeled leucine in arterial plasma. SAIF results were compared with those estimated voxel-wise with the Basis Function Method (BFM) developed by Tomasi and colleagues (Tomasi et al., 2009b), representing the current standard method for quantification of L-[1-^{11}C]Leucine PET.

3.2.3 Passband filter definition

As with the high-pass filter of Turkheimer *et al* (1994) (Turkheimer et al., 1994), the choice of upper and lower cutoff frequencies β_L and β_U for SAIF represents incorporation of prior information about the data into the algorithm. The particular choice of interval influences the final estimated spectrum, and therefore also estimates of the parameters of interest. Since the distributions of the kinetic components in L-[1-^{11}C]Leucine PET data was *a priori* unknown it was impossible to directly define the correspondent passband interval for SAIF application. Thus we set up

a simulation approach based on the information available in literature (Bishu et al., 2008; Tomasi et al., 2009b). The procedure consisted of simulating noisy TACs that were subsequently analyzed by SAIF with different passband intervals. The optimal passband was chosen as the one that provided the best trade-off between precision and accuracy of the estimates of $rCPS$ and other parameters of interest.

Noise-free voxel TACs were simulated according to a heterogeneous kinetic model, with two equilibrating components and one trapping component. Model parameters were generated by the random sampling of normal distributions as follow: V_b=0.064±0.009 (unitless), α_0=0.012±0.002 (ml/cm^3/min), α_1=0.0096±0.0004 (ml/cm^3/min), α_2=0.0197±0.0015 (ml/cm^3/min), β_0=0 (min^{-1}), β_1=0.032±0.004 and β_2=0.15±0.03 (min^{-1}). These values were extracted from (Bishu et al., 2008, 2009) considering tissues with fast and slow kinetics. One thousand noisy TACs were simulated by adding to the noise-free simulated TACs Gaussian noise with zero-mean and variance consistent with the level of noise in ROI and in voxel data. For each noise level, different passband intervals (β_L = 0.005:0.06 min^{-1}; β_U = 0.1:0.6 min^{-1}) were tested. To select the best interval, performance indices percent bias ($\%bias$) and percent root mean square error ($\%RMSE$) were computed. Any noise realization that resulted in a negative entry in the simulated TAC was discarded and replaced with a new realization. Passband filter analysis was performed separately for region and voxel noise level.

At the ROI level, the interval [0.03 min^{-1}; 0.3 min^{-1}] was selected as a good compromise compromised between precision and accuracy, with particular attention to the rCPS parameter. Based on the same criteria, for voxel-wise SAIF quantification we chose 0.02 min^{-1} and 0.2 min^{-1} for β_L and β_U respectively. Due to the large discrepancy in signal-to-noise ration it was expected that the optimal voxel-SAIF passband filter differed somewhat from that used for ROI-SAIF. In particular, high frequency noise present in voxel data but attenuated in the ROI data reduces the accuracy with which high-frequency components can be estimated. Thus, to reduce the impact of high frequency noise, the upper-bound for

the voxel-SAIF passband filter has to be more selective compared to the filter for ROI-SAIF (0.2 min^{-1} instead of 0.3 min^{-1}). At the other end of the spectrum, the less restrictive lower bound of the voxel-SAIF passband (0.02 min^{-1} for voxel-SAIF; 0.03 min^{-1} for ROI-SAIF) has the effect of allowing greater kinetic variability in low-frequency components among individual brain voxels.

3.2.4 Method implementation

Once the choice of passband was made, SAIF was applied to the measured datasets. The distribution of the βs grid as well as the weighting procedure were applied as described in Section 2.3.2. The error variance was modelled assuming Poisson statistics as

$$var\left(C_T(t_i)\right) = \sigma^2 \frac{e^{\gamma t_i} \cdot C_T(t_i)}{\Delta t_i} \tag{3.5}$$

(Wu and Carson, 2002) where γ is the decay constant for ^{11}C, Δt_i is length of Frame i, and σ^2 is a proportionality coefficient reflecting the noise level in the data.

The difference between tracer arrival time in brain and arterial sampling site was estimated by shifting blood curves $0 - 20$ sec, fitting the whole brain TAC, and selecting the delay that produced the smallest WRSS. Tracer appearance times in various parts of the brain differ from the mean of the brain as a whole by ± 2 sec (Iida et al., 1988); therefore, in each study, for all ROIs and all voxels, whole brain tracer arrival delay was used.

For consistency the same setting was extended to BFM also. This decision was taken, in order to avoid differences between the methods related to discrepancy in the algorithm setting. Since the similarity of the methods (Tomasi et al., 2009b; Veronese et al., 2010), the utilization of the a common setting did not represent a critical point for BFM application.

3.2.5 RESULTS

REGION-WISE QUANTIFICATION For what concerns the regional analysis, SAIF produced good fits of ROI TACs (Fig. 3.2.1). In all ROIs one trapping component, one blood component, and at least two equilibrating components were detected. Components detected in frontal cortex, thalamus, and corona radiata in one subject are illustrated in Figure 3.2.1A-C. Of the 170 ROIs analyzed (19 in 8 subjects and 18 in one subject), in 24 ROIs (14%) three equilibrating components were necessary to describe total activity in the region. It was not consistent among subjects, however, which ROIs required the larger number of components. Figure 3.2.1D-F compares the best-fit of measured total activity determined with SAIF to that obtained with voxel-wise BFM analysis in one subject. In general, total activity estimated with SAIF fits measured data as well as or slightly better than the mean of voxel-wise-estimated activities. The two methods also provided comparable trends for predicted time courses of labeled free leucine and protein in tissue. After the first approximately 15-20 min, labeled free leucine tended to be slightly lower, and labeled protein slightly higher, when estimated with SAIF compared with voxel-wise analysis. In the initial portion of the curves, particularly in corona radiata, these trends were reversed. Overall, similarities between estimation methods in predicted time courses of labeled free leucine and protein are reflected in similarities in estimated rCPS in this subject: the percent differences between SAIF and voxel-wise estimates were 1.5%, 1.9%, and 3.1% for frontal cortex, thalamus and corona radiata, respectively.
Figure 3.2.2 illustrates the comparison of region-SAIF estimates with regional mean values of voxel-BFM estimates of K_1, λ, and $rCPS$ in individual subjects. The percent difference between SAIF and voxel-wise estimates for each parameter was computed for each subject. Graphs represent inter-subject means and SDs of the percent differences between analysis methods. In general, differences between methods were small, but inter-subject variability tended to be greater in smaller, and therefore noisier, regions. Across all regions and subjects, the percent difference

in K_1 ranged from a minimum of -12% (in the precentral gyrus of one subject) to a maximum of 6% (hipppocampus); inter-subject mean values ranged from -5% to 1%). The percent difference in , across all regions and subjects, ranged from -14% (hippocampus) to 6% (hypothalamus); inter-subject mean value range: -5% to 4%. The range of percent differences in rCPS, across all regions and subjects, was -18% (occipital cortex of one subject) to 23% (hippocampus); inter-subject mean values ranged from 9% to 7%.

Table 3.2.1 summarizes values of K_1, λ, and $rCPS$ estimated by ROI--SAIF and those estimated voxel-wise with BFM. Means and SD for 9 subjects are reported for 18 ROIs and whole brain. There is good agreement between the mean values of the two methods in most regions, but inter-subject variability was somewhat higher in λ and $rCPS$ estimated with SAIF than with the BFM method. In most regions, there was also good agreement between the methods in estimates of V_b. SAIF and mean voxel-wise estimates of V_b in whole brain were 0.062$\pm$0.009 and 0.062$\pm$0.008, respectively. In all regions examined, the inter-subject mean of the SAIF estimates of V_b was 95% to 102% the corresponding mean of the voxel-wise estimates.

VOXEL-WISE QUANTIFICATION When SAIF was applied at voxel level, a superior agreement was found with BFM method in the description of all the parameters of interest. Respect to regional-SAIF, this behavior was awaited, since the two methodologies were compared at the same resolution level.

Table 3.2.2 summarizes the relative difference of mean voxel estimates within ROIs of K_1, λ, and $rCPS$. Results refer to those obtained with SAIF and BFM across subjects. The relative $rCPS$ difference across ROIs and subjects was 6%$\pm$5% (mean$\pm$SD). The best agreement was found in Thalamus and Putamen (1.5% and -1.9% respectively), while the worst was detected in Hippocampus and Amigdala (17% and 15%, respectively).

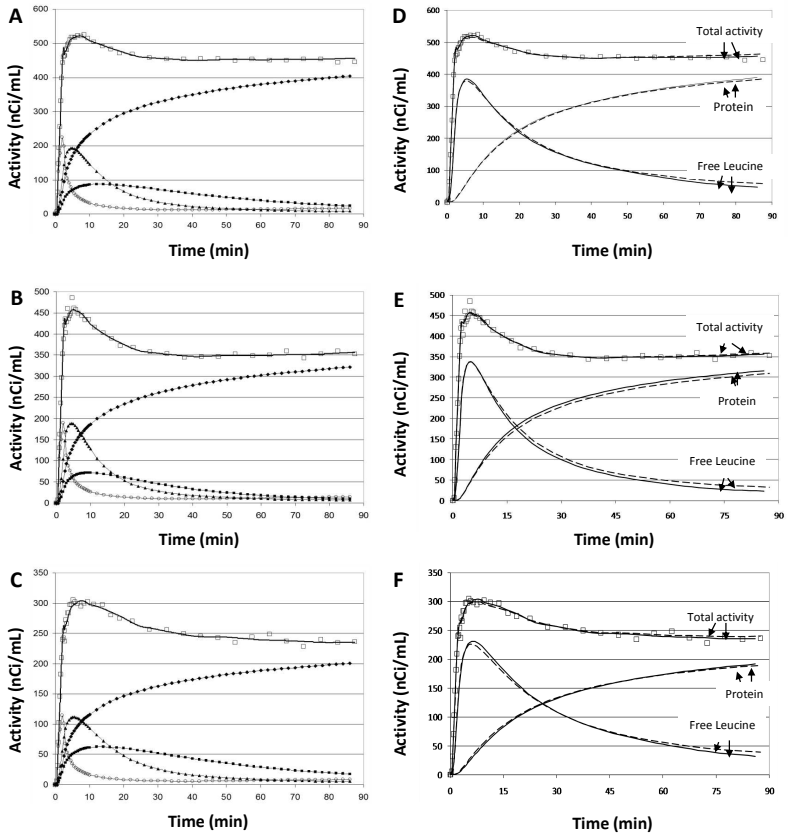

Figure 3.2.1: Time courses of total activity and activity in tissue components estimated with SAIF. Total activity was measured in frontal cortex (A), thalamus (B) and corona radiata (C) in one subject. Open squares and solid lines represent the measured and fitted total activity in the ROIs. Open circles represent activity of blood in the brain. Solid diamonds, solid squares and solid triangles represent trapping, slow, and fast components, respectively, detected by SAIF. Concentration of $^{11}CO_2$ in brain is not shown. (D-F) SAIF and voxel-wise fits. The graphs show the time courses of total activity, labeled protein and labeled free leucine in frontal cortex (D), thalamus (E) and corona radiata (F) of one subject. Open squares indicate measured total activity in the ROI, and solid lines represent activity estimated by fitting the ROI TAC using SAIF. The average time course of the activities determined in each voxel from the parameter estimates for that voxel is shown by the dashed lines. Free leucine and protein time courses estimated by the two methods are similar in the very early portion of the curve; they gradually separate with increasing time. In the last 15-30 min, one sees slightly lower total activity and free leucine estimated by SAIF, and slightly higher concentrations of labeled protein.

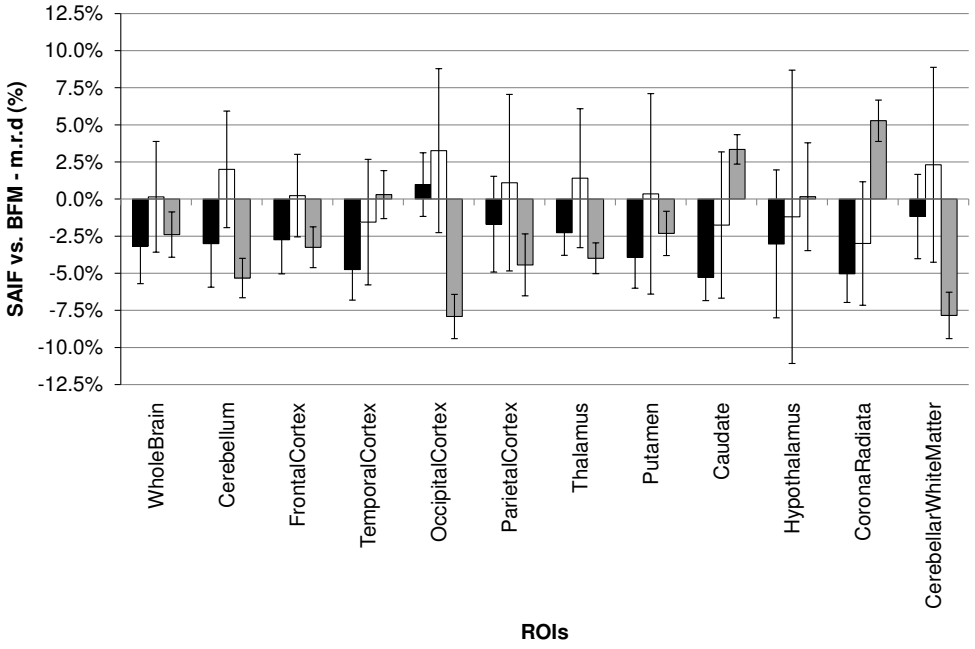

Figure 3.2.2: Percent difference between SAIF and voxel-wise estimates of K_1 (A), λ (B) and $rCPS$ (C). For each of the nine subjects, for each ROI, and for each parameter, the percent difference was computed as $100 * (p_{SAIF} - p_{BFM})/p_{BFM}$, where p_{SAIF} is the SAIF estimate and p_{BFM} is the mean of the voxel-wise estimates of the parameter. For each ROI indicated, the bar represents the inter-subject mean of the percent difference in the parameter estimates. The inter-subject standard deviation is indicated by the error bar.

Table 3.2.1: Regional Estimates of K_1, λ and $rCPS$

Region	ROI TAC-based Estimation with SAIF			Voxelwise Estimation with BFM		
	K_1 $(mL/g/min)$	λ (unitless)	$rCPS$ $(mL/g/min)$	K_1 $(mL/g/min)$	λ (unitless)	$rCPS$ $(mL/g/min)$
Whole Brain (1,397±133 cm^3)	0.047 ± 0.004	0.77 ± 0.03	1.57 ± 0.13	0.049 ± 0.004	0.77 ± 0.02	1.61 ± 0.08
Cerebellum (115±15 cm^3)	0.067 ± 0.007	0.80 ± 0.02	1.85 ± 0.14	0.07 ± 0.007	0.78 ± 0.02	1.96 ± 0.10
Vermis (6±1 cm^3)	0.064 ± 0.007	0.79 ± 0.04	1.85 ± 0.16	0.066 ± 0.008	0.77 ± 0.03	1.96 ± 0.09
CORTICAL						
Frontal cortex (135±61 cm^3)	0.051 ± 0.005	0.75 ± 0.03	1.83 ± 0.11	0.052 ± 0.005	0.75 ± 0.02	1.89 ± 0.09
Temporal cortex (58±12 cm^3)	0.044 ± 0.004	0.73 ± 0.03	1.74 ± 0.15	0.046 ± 0.004	0.75 ± 0.02	1.73 ± 0.10
Occipical cortex (29±12 cm^3)	0.064 ± 0.006	0.78 ± 0.03	1.93 ± 0.21	0.063 ± 0.006	0.76 ± 0.03	2.09 ± 0.17
Parietal cortex (17±5 cm^3)	0.053 ± 0.004	0.76 ± 0.04	1.78 ± 0.28	0.054 ± 0.004	0.76 ± 0.02	1.86 ± 0.2
Supraorbital gyrus (4±1 cm^3)	0.047 ± 0.006	0.74 ± 0.03	1.76 ± 0.14	0.048 ± 0.005	0.75 ± 0.03	1.77 ± 0.10
Precentral gyrus (3±0 cm^3)	0.057 ± 0.005	0.78 ± 0.04	1.75 ± 0.29	0.056 ± 0.006	0.76 ± 0.02	1.88 ± 0.22
Postcentral gyrus (2±0 cm^3)	0.059 ± 0.006	0.79 ± 0.03	1.71 ± 0.16	0.059 ± 0.006	0.77 ± 0.02	1.87 ± 0.18
SUBCORTICAL						
Thalamus (15±6 cm^3)	0.05 ± 0.004	0.79 ± 0.02	1.47 ± 0.07	0.051 ± 0.003	0.78 ± 0.02	1.54 ± 0.10
Caudate (8±1 cm^3)	0.033 ± 0.004	0.77 ± 0.02	1.04 ± 0.10	0.034 ± 0.005	0.79 ± 0.02	1.01 ± 0.07
Putamen (9±1 cm^3)	0.048 ± 0.005	0.80 ± 0.03	1.3 ± 0.13	0.051 ± 0.006	0.80 ± 0.02	1.33 ± 0.08
Hippocampus (4±1 cm^3)	0.035 ± 0.004	0.71 ± 0.04	1.58 ± 0.2	0.037 ± 0.004	0.74 ± 0.02	1.47 ± 0.15
Amygdala (5±1 cm^3)	0.031 ± 0.003	0.71 ± 0.05	1.42 ± 0.18	0.033 ± 0.003	0.74 ± 0.03	1.35 ± 0.11
Hypothalamus (0.6±0.2 cm^3)	0.035 ± 0.015	0.69 ± 0.26	1.11 ± 0.53	0.041 ± 0.005	0.79 ± 0.03	1.24 ± 0.24
WHITE MATTER						
Corona radiata (46±5 cm^3)	0.026 ± 0.004	0.76 ± 0.02	0.87 ± 0.08	0.027 ± 0.004	0.79 ± 0.02	0.83 ± 0.06
Centrum semiovale (35±11 cm^3)	0.028 ± 0.003	0.76 ± 0.03	0.97 ± 0.12	0.03 ± 0.004	0.78 ± 0.02	0.91 ± 0.08
Cerebellar peduncles (14±2 cm^3)	0.043 ± 0.004	0.81 ± 0.03	1.07 ± 0.13	0.044 ± 0.004	0.80 ± 0.02	1.16 ± 0.08

Similarly to $rCPS$, limited discrepancy was found also for K_1 (mean±SD relative difference = -4%±4%) and λ (mean±SD relative difference = -3%±2%.

The same good agreement between SAIF and BFM can be noticed in Figure 3.2.3. The scatter analysis of the mean of voxel estimates within regions showed Perason's correlation coefficient (R^2) ranged from 0.89 (for λ) (Fig. 3.2.3A) up to 0.99 (for $rCPS$) (Fig. 3.2.3B). No statistical difference (p <0.05) was found by the comparison of mean of voxel estimates between SAIF and BFM (Fig. 3.2.3C).

Figure 3.2.4 allows a visual comparison of $rCPS$ and λ parametric maps, obtained with BFM and SAIF respectively. Coherently with previous results, the spatial distribution of the voxel estimates appears very similar between different methods.

In term of fit of the data (Fig. 3.2.5) SAIF and BFM provided similar description. Altough WRSS was generally lower with SAIF, the difference with BFM was $< 1\%$ (mean across all the subjects). The total rate

Table 3.2.2: Percent relative differences of voxel-wise SAIF and BFM for K_1, λ, and $rCPS$

Region	K_1			λ			$rCPS$		
WholeBrain	-4.90%	±	1.90%	-3.10%	±	1.40%	5.10%	±	3.60%
Cerebellum	-11.00%	±	2.10%	-4.30%	±	0.80%	6.50%	±	3.90%
CORTICAL									
Frontal cortex	-3.00%	±	2.90%	-2.90%	±	1.90%	4.40%	±	3.80%
Temporal cortex	-1.80%	±	1.80%	-2.80%	±	0.90%	4.70%	±	4.10%
Occipital cortex	-6.20%	±	2.00%	-2.30%	±	1.30%	2.30%	±	2.80%
Parietal cortex	-6.10%	±	3.20%	-3.50%	±	2.30%	4.60%	±	3.80%
SUBCORTICAL									
Thalamus	-5.80%	±	2.70%	-1.80%	±	0.90%	1.50%	±	4.20%
Putamen	-0.30%	±	2.40%	0.00%	±	0.90%	-1.90%	±	5.00%
Caudate	0.40%	±	3.00%	-2.00%	±	2.00%	4.20%	±	8.30%
Hippocampus	-1.10%	±	3.30%	-6.60%	±	2.60%	16.60%	±	6.60%
Amygdala	-2.10%	±	3.00%	-6.60%	±	2.10%	14.70%	±	5.00%
WHITE MATTER									
Corona Radiata	0.70%	±	4.10%	-2.40%	±	2.40%	7.10%	±	8.30%
Cerebellar peduncles	-8.00%	±	3.00%	-2.90%	±	1.20%	4.90%	±	5.20%
Mean	-3.80%			-3.20%			5.80%		
SD	3.50%			1.80%			5.00%		

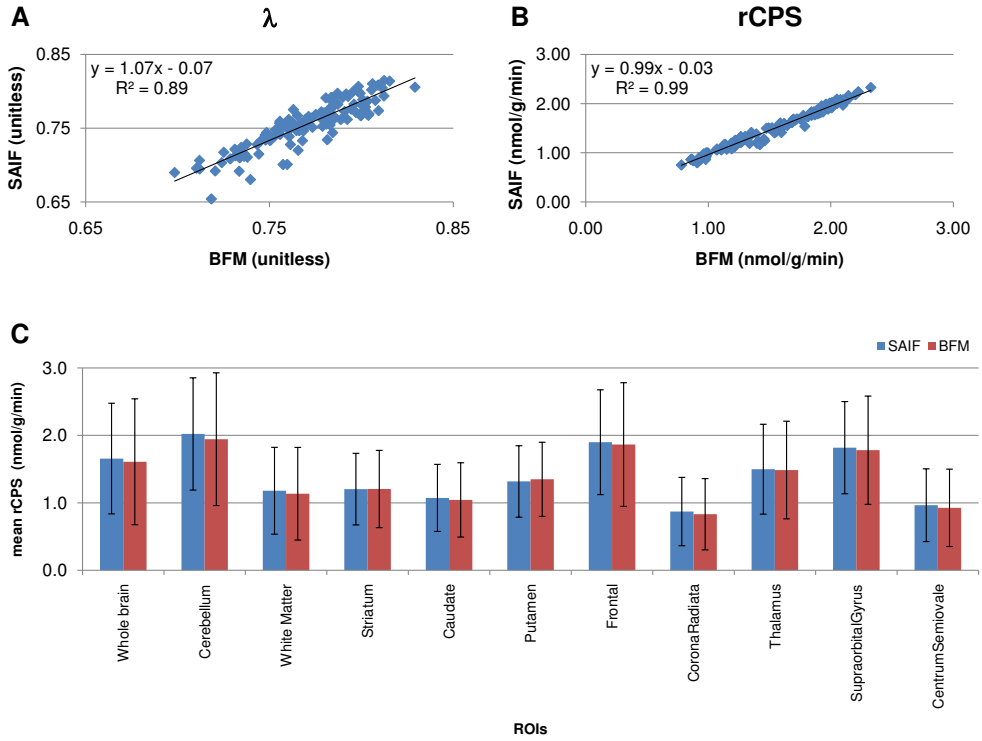

Figure 3.2.3: BFM and SAIF comparison.

of failures and outliers was low with both BFM (0.4%±0.3%) and SAIF (1.7%±1.0%). The fraction of voxels with zero rCPS was almost negligible (<1%) for all the methods and all the tested scenarios. In term of number and type of components return with voxel-SAIF, one trapping and one blood component were found in 100% of successfully estimated voxels, but, differently ffrom regional analysis, only 45% of them included two or more equilibrating components. The finding of more than one equilibrating component suggests kinetic heterogeneity of the tissue in these voxels. In the remaining voxels only one equilibrating component was detected, indicating the homogeneous tissue model was appropriate to describe 55% of brain voxel TACs.

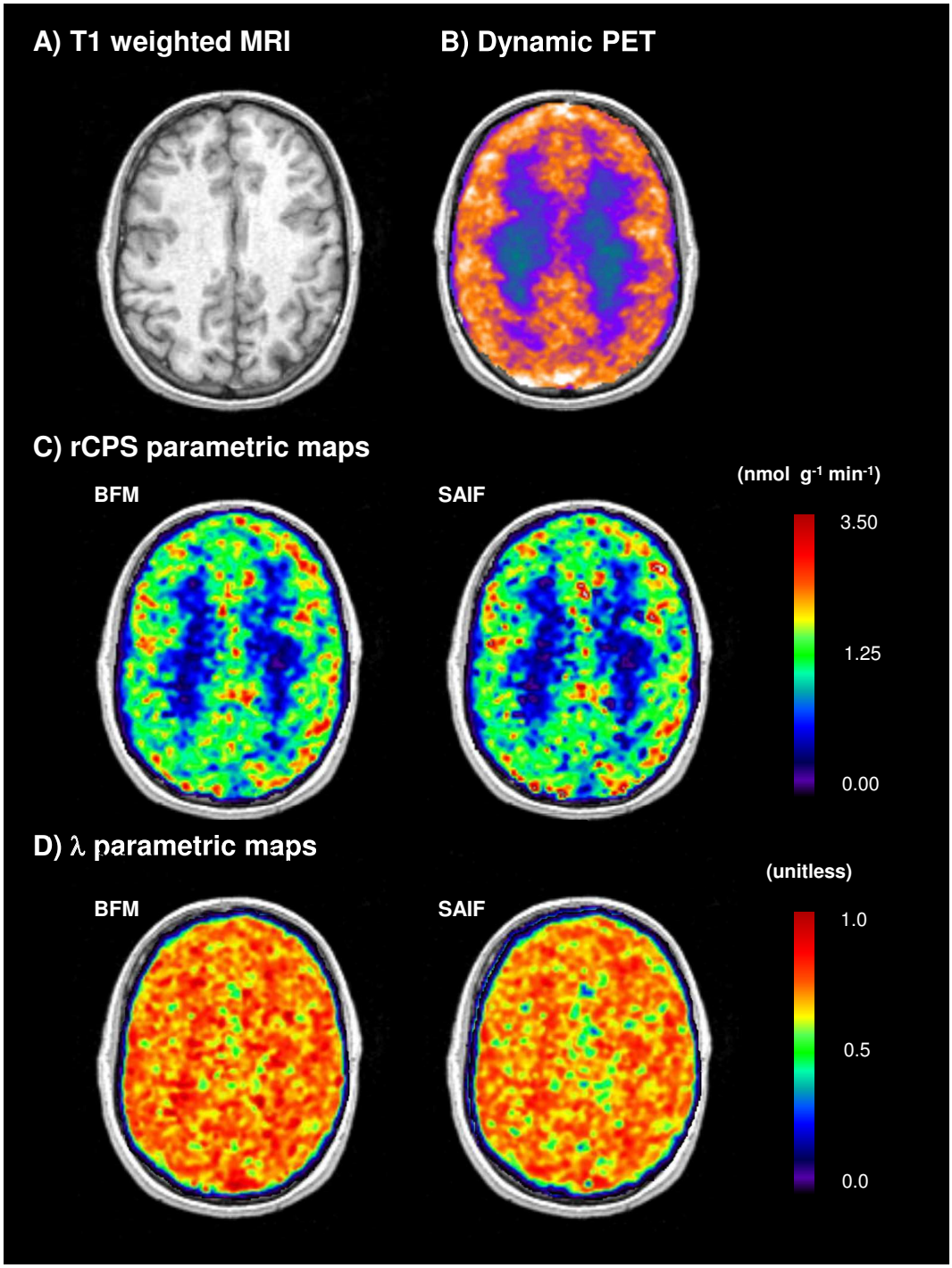

Figure 3.2.4: Parametric Maps.

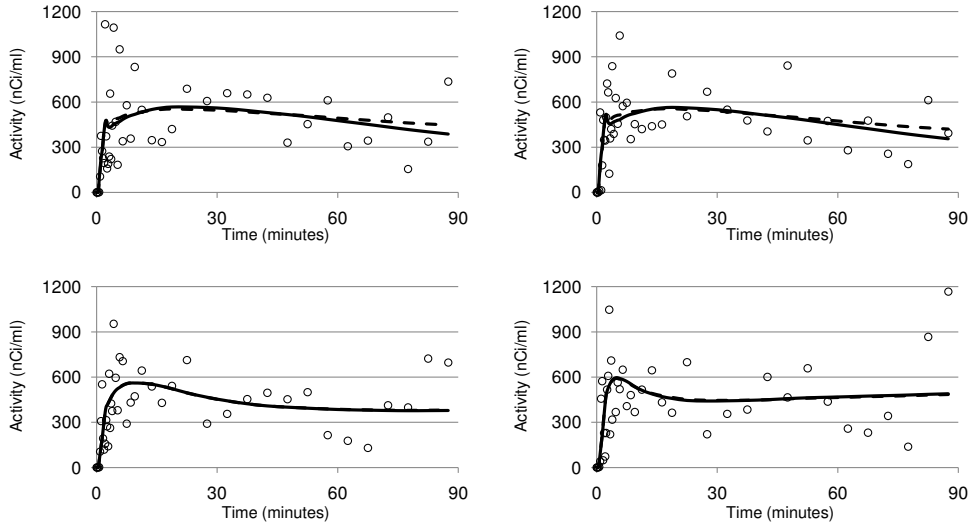

Figure 3.2.5: BFM and SAIF model fit of the data.

3.2.6 CONSIDERATIONS

Use of the iterative filter implemented in SAIF improved the robustness of the method to noise; SAIF provided good estimates of leucine kinetic parameters and $rCPS$, for both low and high noise level. ROI-based- and mean voxel-based $rCPS$ estimates were highly correlated with BFM, the current state of the art method for leucine kinetic quantification.
Very interesting considerations can be done when means of voxel-SAIF estimates within a region are compared to ROI-SAIF estimates. In general, even though two different levels of resolution (ROI and voxel) were compared each other, it can be found good agreement between analysis methods. Unlike fixed compartmental model analyses, spectral analysis allows the tissue ROI to be represented as a linear combination of any number of tissue compartments. Because the tissue ROI could be represented as a linear combination of the compartments of each voxel comprising the ROI, we could expect good agreement between the ROI-SAIF estimates and mean voxel-SAIF estimates of linear parameters of the model. Indeed, the linear parameter V_b and the parameter K_1, which

is the sum of the linear coefficients α_i, exhibit the best agreement (1%±4% and 1%±4%, mean±SD of the difference between ROI-SAIF and mean voxel-SAIF estimates of V_b and K_1, respectively) and highest correlation between methods (R^2 = 0.97 for both V_b and K_1). In practice, of course, the number of components identified in ROI data is far smaller than the sum of the unique components found in the voxels of the ROI, principally due to the small number of observations in a single ROI TAC, which limits the number of components that can be identified. In contrast to linear parameters, λ and $rCPS$ are non-linear combinations of parameters, and therefore less agreement between ROI-SAIF and mean voxel-SAIF estimates is to be expected. Estimates of λ were 2%±2% higher with ROI-SAIF and correlation between estimates provided by the two methods was 0.67. Correlation between $rCPS$ estimated with the two methods was very high (R^2=0.92), but estimates were 8%±6% lower with ROI-SAIF. Determination of λ from the SAIF-estimated coefficients α_i is based on the assumption that all λ in the subregions of a heterogeneous tissue are equal; this constraint is likely to have a greater impact when data are analyzed at the ROI level, where a uniform value of λ for the entire region is assumed, than when data are analyzed at the voxel level and λ can vary among voxels within the region. The constraint contributes to differences between ROI-SAIF and voxel-SAIF estimates of λ and, estimates of $rCPS$. For this reason, voxel-SAIF method should be preferred from the ROI one.

Essential to performance of SAIF is selection of appropriate upper and lower bounds of the passband. The passband interval can be thought of as prior information on the data; it is defined as the interval in which true equilibrating components in the system are expected to be found. We chose passband limits based on results of a simulation study, but if the true kinetics of leucine in the tissue differ appreciably from those used in the simulation, the optimal passband may differ from the one we selected. Other approaches to selection of passband endpoints may lead to further improvements in SAIF performance and should be explored.

3.3 [^{11}C]SCH442416 PET STUDIES IN HEALTHY SUBJECTS

3.3.1 INTRODUCTION

A2A receptor represents one of the four types of adenosine receptors. A2A plays roles in the heart, regulating myocardial oxygen consumption and coronary blood flow, and it also has broader anti-inflammatory effects throughout the body (Hasko and Pacher, 2008). In the brain A2A are involved in the regulation of the neurotransmitter release such as dopamine and glutamate, making it a potential therapeutic target for the treatment of conditions such as insomnia, pain, depression, drug addiction and Parkinson's disease (Cunha et al., 2008; Fuxe et al., 2007; Mihara et al., 2008; Schiffmann et al., 2007).
[^{11}C]SCH442416 is a nonxanthine radioligand which binds selectively and reversibly to striatal A2A receptors, abundant in basal ganglia, vasculature and platelets (Moresco et al., 2005). A2A receptor interacts structurally and functionally with the dopamine D2 receptor and this interaction is of particular interest as it is thought to be central to basal ganglia dysfunction in Parkinson's disease (Schiffmann et al., 2007). Preclinical studies on rats and nonhuman primates suggest that [^{11}C]SCH442416 is suitable for the in vivo imaging of adenosine A2A receptors with PET because of its high affinity and selectivity, good signal-to-noise ratio, and low levels of radioactive metabolites in the brains (Moresco et al., 2005; Todde et al., 2000).
Quantitative assessment of A2A receptor expression from [^{11}C]SCH442416 PET images in human brain have been made by using spectral analysis (Hinz et al., 2003) revealing the presence of both reversible and irreversible components. From SA application it has been possible to define [^{11}C]SCH442416 compartmental model (Rizzo, 2012) (Fig. 3.3.1), composed of one reversible tissue compartment and one additional irreversible compartment. The latter is required to describe the vascular kinetics of the tracer.
These conditions might appear not adequate for SAIF applicability, which

has been defined for tracers with irreversible tissue kinetics. However, the flexibility of the SA model could theoretically allow to extend the use of SAIF even in presence of vascular trapping. The aim of this section is to verified this possibility.

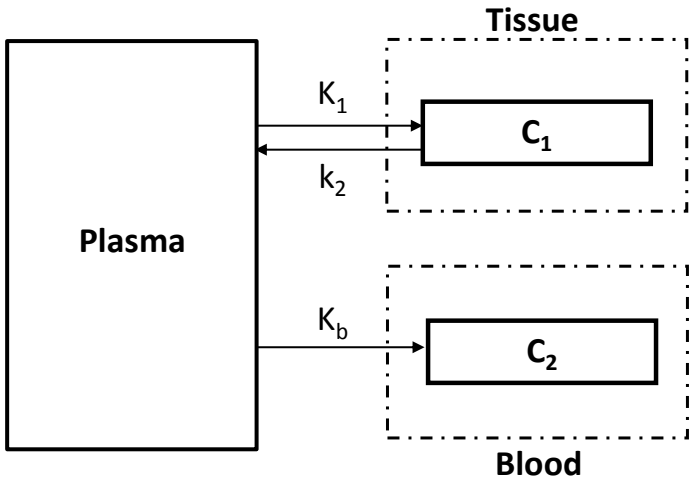

Figure 3.3.1: [^{11}C]SCH442416 compartmental model. Two tissue-three rate constants in parallel ($3K$-parallel) compartmental model. K_1 (ml/cm^3/min) and k_2 (min^{-1}) represent first order rate constants for transport of ligand from plasma to tissue (C_1) and vice versa. K_b (ml/cm^3/min) represents the trapping vascular component (C_2).

3.3.2 Adaptability of SAIF to tracer with blood trapping

As indicated by Rizzo (2012), the correct model for describing the [^{11}C]SCH442416 kinetics consists in:

$$\begin{cases} \frac{C_1(t)}{dt} &= K_1 C_p(t) - k_2 C_1(t) & C_1(0) = 0 \\ \frac{C_2(t)}{dt} &= K_b C_p(t) & C_2(0) = 0 \\ C_{measured}(t) &= (1 - V_b) C_1(t) + V_b \left[C_2(t) + C_b(t) \right] & \end{cases} \quad (3.6)$$

where K_1 (ml/cm^3/min) and k_2 (min^{-1}) are the first order rate constants for transport of ligand from plasma to tissue and vice versa. K_b (ml/cm^3/min) represents the rate constant from plasma to the trapping vascular component C_2.
From Eq. 3.6, the explicit solution can be derived as

$$C_{measured}(t) = (1 - V_b)\left[K_1 C_p(t) \otimes e^{-k_2(t)}\right] + V_b\left[K_b \int_0^t C_p(\tau)d\tau + C_b(t)\right] \tag{3.7}$$

From the comparison between Eq. 3.7 and SA model equation (Eq. 2.7), a close relationship can be find. The only difference is represented by the trapping component that in standard SA is related to the tissues (and for this reason multiply by the term $(1 - V_b)$) while in [^{11}C]SCH442416 is related to the blood (and for this reason multiply by the term V_b).
Thus, it is not difficult to modify the SA equation for both ESA and SAIF in order to extend the method applicability to [^{11}C]SCH442416 data. Once the kinetic spectra are estimated, it is sufficient to rescale the trapping coefficient for the blood volume fraction instead of the tissue fraction. Moreover, since the K_b indicated the trapping in the vascular compartment, all the equilibrating components found with SA model can be assumed as belonging to the kinetic of the tracer in tissues and therefore used to compute the distribution volume. Thus, the same formula of Eq. 2.5 here recall

$$V_T = \sum_{j=1}^{M} \frac{\alpha_j}{\beta_j} \tag{3.8}$$

3.3.3 Dataset and methods of analysis

Data of [^{11}C]SCH442416 of 5 healthy subjects from previous study were used (Hinz et al., 2003). Each subject underwent a 90-min dynamic PET scanning in a Siemens ECAT EXACT HRD scanner after a bolus injection of an average of 612 MBq of [^{11}C]SCH442416 over 10 seconds, 30 seconds after the start of the scan. All PET data were acquired in 3-dimensional mode, corrected for attenuation, detector efficiency, random

events, and scatter, and reconstructed into tomographic images using filtered back-projection[9]. Acquisition was performed in list-mode (event by event) and scans were rebinned into 34 time frames (6x10, 3x20, 3x30, 4x60, 6x120, 8x300, 3x600 seconds). The reconstructed voxel sizes were 2.096 mm x 2.096 mm x 2.43 mm. After injection of [^{11}C]SCH442416 the radioactivity concentration in blood was measured continuously for the first 15 minutes with discrete blood samples taken at baseline, 5, 10, 15, 20, 30, 40, 50, 60, 75, and 90 minutes. Eight of these samples were also used for quantification of the fraction of radioactivity attributable to unmetabolized parent radiotracer, generating the metabolite-corrected arterial plasma input function for all subjects. The tissue, blood and plasma data were corrected for decay of ^{11}C.
Each subject had a T1-weighted MRI used only for region of interest placement through the coregistration with a maximum probability brain atlas (Hammers et al., 2003). Among all the 73 regions present in the atlas, Hippocampus , Amygdala, Cerebellum, Caudate, Putamen, Thalamus, Temporal Cortex, Pallidum and Frontal Cortex were selected for the analysis. Right and left hemispheres were considered combined.
As current state of the art for [^{11}C]SCH442416 quantification we selected:

- Weighted Non Linear Least Squares (WNLLS) applied to literature reference [^{11}C]SCH442416 compartmental model (Fig. 3.3.1);
- Standard and not filtered SA (ESA);

All the methods were applied voxel-wise and results were compared with those provided by SAIF applied at the same resolution level. For each subject, SAIF passband was derived by WNLLS results, by setting the filter bounds equal to the minimum and maximum of k_2 parameter estimated with the non linear method. This criterion was defined by considering the relationship between [^{11}C]SCH442416 (Eq. 3.7) and SA model.
Distribution volume of the tracer in tissue, V_T (ml/cm^3) was considered as parameter of interest. With WNLLS, V_T was calculated as the ratio of tissue kinetic rate constant ($V_T = K_1/k_2$). Mean of voxel estimates within

regions and outliers were used for the computation of the performance indexes.

3.3.4 RESULTS

Figure 3.3.2 reports a comparison of mean of voxel V_T estimates within ROIs, by considering all the tested methods. The same subset of voxels in which all the methods successfully converged to physiological solutions were considered. While WNLLS and ESA correlation was clearly weak (R^2=0.55), SAIF and WNLLS estimates resulted well correlated (R^2=0.91) (Fig. 3.3.2A). Also the inter-subject variability of V_T estimates were comparable between WNLLS and SAIF, while with ESA the variability between subjects resulted higher (Fig. 3.3.2B). The same relationship between the methods can be noticed by considering the V_T parametric maps (Fig. 3.3.3). The spatial distribution of V_T values computed with ESA is clearly not coherent with those reported by SAIF and WNLLS. Even if we do not know the true distribution of these values, due to the physiological organization of the brain tissues and their spatial smoothness, we can hypothesize that the results provided with ESA are unreliable respect to those provided by SAIF and ESA.
Despite this difference, all the methods provided almost the same description of the data (Fig. 3.3.4): WRSS and RSS derived from WNLLS application were found approximately 10% higher respected to those returned by ESA and SAIF.

3.3.5 CONSIDERATIONS

[^{11}C]SCH442416 kinetics showed an important vascular component whose effect was to complicate the anatomical atlas co-registration to the PET images as well as to mislead the estimation of the model parameters.
This characteristic combined with the high sensitivity of the noise can be the justification for which ESA method did not provide reliable results.

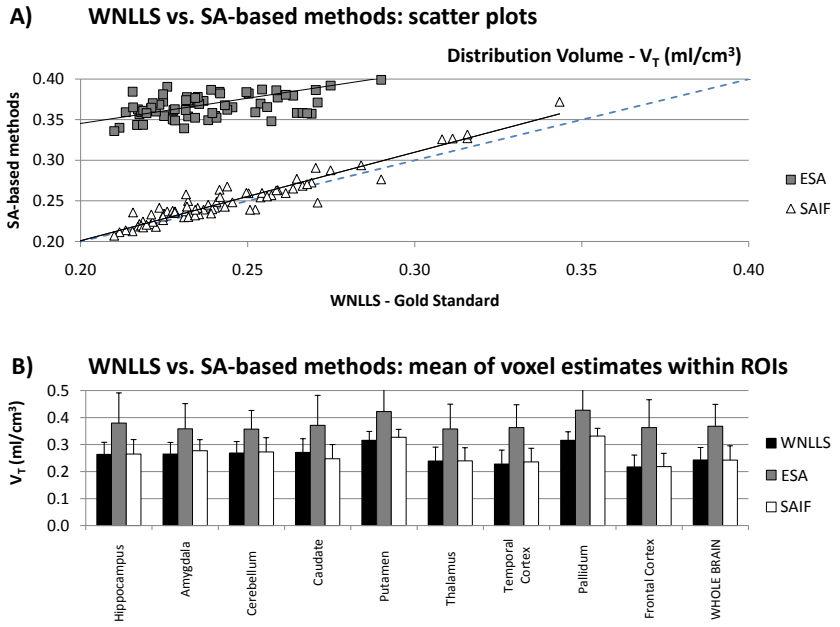

Figure 3.3.2: WNLLS vs. SA-based methods. Panel A shows the scatter plot of mean of V_T estimates within ROIs, calculated with WNLLS (x axis) and SA-based methods (y axis). All the subjects and ROIs are reported. Open triangles refer to SAIF. Gray squares refer to ESA. Panal B compares the mean of voxel V_T estimates within ROIs. SD represents inter-subject variability. Black bars refer to WNLLS, gray bars refer to ESA and white bars refer to SAIF.

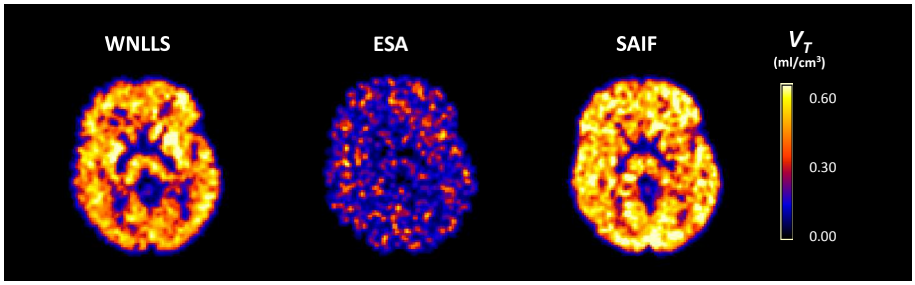

Figure 3.3.3: Comparison of V_T parametric maps.

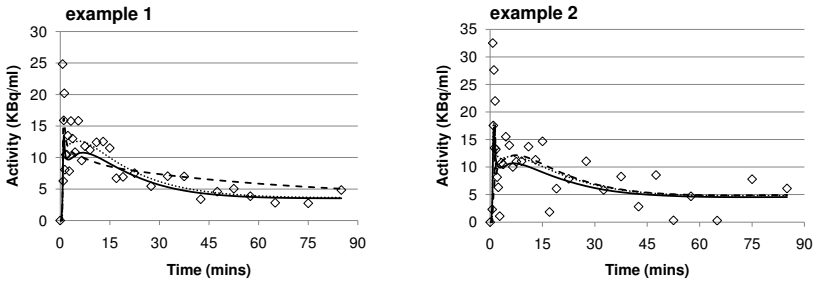

Figure 3.3.4: Data fit performance. Solid lines refer to WNLLS, dotted line refer to ESA and dashed line refer to SAIF.

By considering the results presented in literature (Galazzo et al., 2010; Rizzo, 2012) we can suggest to limit ESA application to [^{11}C]SCH442416 data only for regional analysis. On the contrary, SAIF performed similarly to WNLLS method. The good agreement was verified in term of parametric maps and description of the data. Respect to the non linear method, the advantage of SAIF can be represented by its computational time (reduced by 50%). However, if we considered the outliers, both SAIF and WNLLS were penalized by a high fraction of not physiological estimates (in some subjects >30% of the whole field of view). This suggests that SAIF as well as WNLLS might not represent the best approaches for quantification of [^{11}C]SCH442416 PET data and that different alternatives, like the HMAP or HBFM developed by Rizzo (Rizzo et al., 2012a,b), should be considered.

If we don't succeed, we run the risk of failure.

Dan Quayle

4

Applications to non brain studies

THIS CHAPTER PRESENTS some representative applications of SAIF in non brain PET studies, with the aim to demonstrate the adaptability of the technique to physiological systems different from brain, the anatomical context for which SAIF was originally developed. SAIF performance will be compared with those provided by the standard and already validated methods of analysis, in order to stress advantages and drawbacks of the presented quantification tool respect to what is currently used in literature.

4.1 [^{18}F]FDG PET STUDIES IN SKELETAL LEG MUSCLE

4.1.1 INTRODUCTION

The evaluation of SAIF performance cannot be complete without an application with [^{18}F]FDG , the most used PET irreversible tracer in clinical practice. In PET imaging [^{18}F]FDG has been widely applied for the assessment of glucose metabolism in the heart (Miyauchi and Wahl, 1996), lungs (Service, 2005) and the brain (Bartenstein et al., 2002). After [^{18}F]FDG injection, the tracer is taken up by cells where it is phosphorylated by hexokinase and hence trapped. This results into two important consequences:

- the phosphorylation prevents the glucose from being released again from the cell, once it has been absorbed. Thus no metabolites product are cycling in the system;
- the spatial distribution of [^{18}F]FDG tracer uptake directly reflects the metabolic activities of the tissues.

This last characteristic is the main reason for why [^{18}F]FDG -PET has been so largely utilized for tumor imaging in oncology (Mihailovic et al., 2012) where it is used for diagnosis, staging, and monitoring treatment of cancers.
In this section we considered [^{18}F]FDG for the investigation of metabolism of skeletal muscle and its relationship with insulin-glucose system. The analysis requires to consider the entire interactions between delivery, transport and phosphorylation of glucose in governing the uptake into human skeletal leg muscle (Bertoldo et al., 2006).
Generally the main methods for quantitative [^{18}F]FDG imaging are SUV and Patlak. Both of them, however, return only the information about the tracer uptake, which is not sufficient when the purpose is to fully-characterized the kinetic of the tracer. An alternative for [^{18}F]FDG quantification is represented by WNLLS applied to [^{18}F]FDG skeletal muscle compartmental model (Bertoldo et al., 2001). WNLLS demonstrated to

be extremely precise and accurate when it is applied to high SNR level (Bertoldo et al., 2001), but at high noise level like in voxel-wise analysis, its applicability results limited. On the contrary, SAIF might represent a valid alternative for the parametric imaging of [18F]FDG in skeletal muscle, by combining robustness to the noise with high informative outcomes.

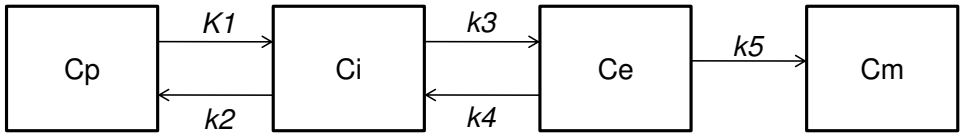

Figure 4.1.1: [18F]FDG compartmental model. K_1 and k_2 refer to reversible exchange between plasma and the interstitial space (Ci); k_3 and k_4 refer to reversible exchange attributed to bidirectional glucose transport (Ce); k_5 describes the kinetic of glucose phosphorylation.

4.1.2 Dataset and methods of analysis

[18F]FDG data of 5 healthy subjects from previous study (Bertoldo et al., 2006) were used. Each subject underwent a 90-min dynamic PET scanning after a bolus injection of an average of 222 MBq (6 mCi) of [18F]FDG over 20 seconds. A Siemens/CTI ECAT HR+ PET scanner in 3-D imaging mode (63 parallel planes, axial field-of-view 15.2 cm, slice width 2.4 mm) was used. The final reconstructed PET image resolution was about 6 mm. Participants were positioned in the PET scanner with the mid-calf area in the center of the field, and to minimize movement during PET imaging, the legs were supported with pliable block molding. MRI was also performed for the ROI placement. For further details the interested reader is referred to (Bertoldo et al., 2006).
As reference methods for [18F]FDG quantification we considered:

- WNLLS applied to [18F]FDG skeletal muscle model. This coincides with a 3-tissue model, composed by two equilibrating compart-

ments (interstitial space and cellular compartments) and one trapping (Fig. 4.1.1) (Bertoldo et al., 2001);

- Patlak graphical analysis;

All the methods were applied voxel-wise and their results were compared with those provided by ESA and SAIF applied at the same resolution level. SAIF passband interval was derived according to previous [^{18}F]FDG study in human skeletal muscle (Bertoldo et al., 2001) and fixed to 0.05 min^{-1} and 0.7 min^{-1} for lower and upper bound respectively. K_i was considered as the main parameters of interest for the evaluation of the method performances. Mean of voxel estimates within regions and outliers were also computed performance indexes.

4.1.3 RESULTS

The comparison of K_i parametric maps obtained for the mid-calf area in one representative subject (Fig. 4.1.2), provides a direct information about the method performances: SAIF parametric maps resulted in the best anatomical description of the investigated tissues. In particular it can be seen how well delineated are the *soleos* and *tibialis* muscles (Panel A), respect to the other methods. In fact, moving from top to the bottom and from left to right (Fig. 4.1.2B-D), there is an evident decrease of method performances coupled with a voxel outlier increase, especially for WNLLS and ESA methods. This is consequence of the different method sensitivity to the noise in the data.
However, when the same subsets of voxels in which all the methods successfully converged into physiological estimates were selected, high correlation (R^2>0.97) and consistency was found for K_i parameter, for all subjects and all regions (Fig. 4.1.3 A). The main difference was related to the number of voxel outliers: SAIF proved to be robust with a voxel failure rate <1%, whereas more than 40% of ESA estimates had to be eliminated because out of the physiological range(Fig. 4.1.3B).
Figure 4.1.4 shows some an additional outcomes of SAIF. In panel A two

representative voxel-TACs (one for muscle and one for low-uptake tissue) are reported. This information is not available with SUV and Patlak methods and, due to the high fraction of failures, it is only partially obtainable with WNLLS. In these reported voxels (Fig. 4.1.4A), for example, WNLLS did not converge to physiological solutions. Panel B and C, instead, show the SAIF parametric maps for K_1 and V_b. As for the model description of the data, this information was not available with SUV and Patlak.

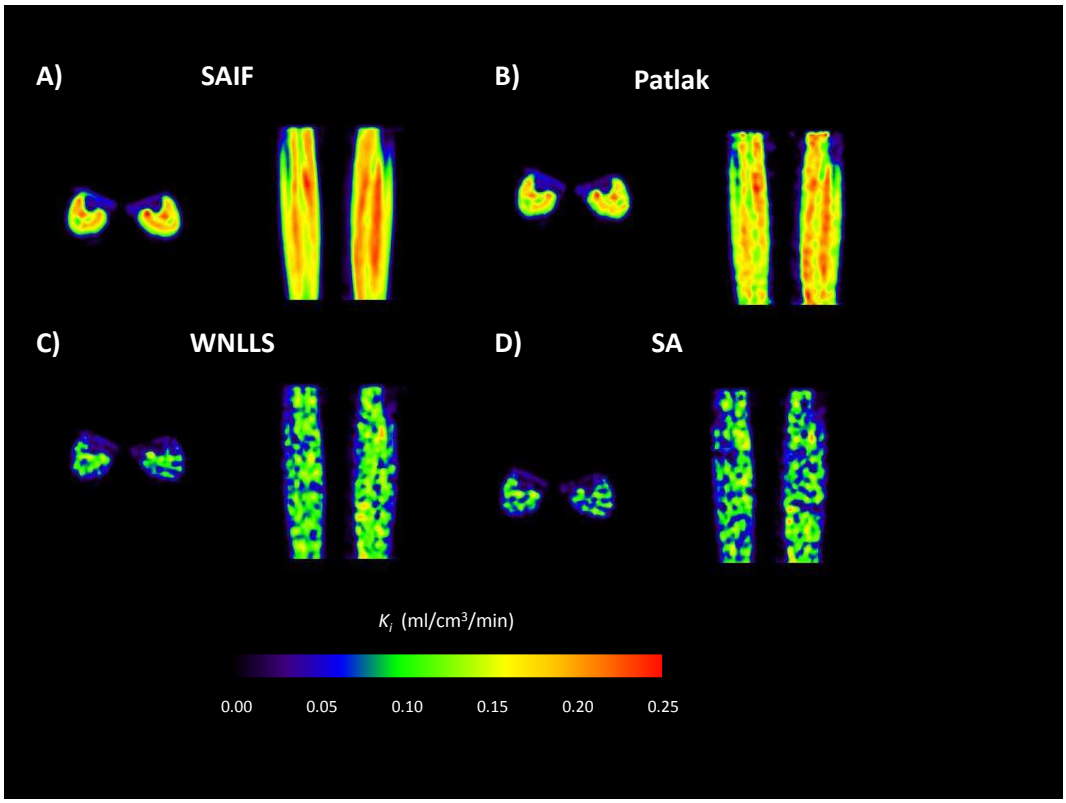

Figure 4.1.2: K_i Parametric maps. Panel A-D show the K_i parametric maps obtained respectively with SAIF, Patlak, WNLLS and ESA. Results refer to the mid-calf area in one representative subject, reported in both axial and coronal views.

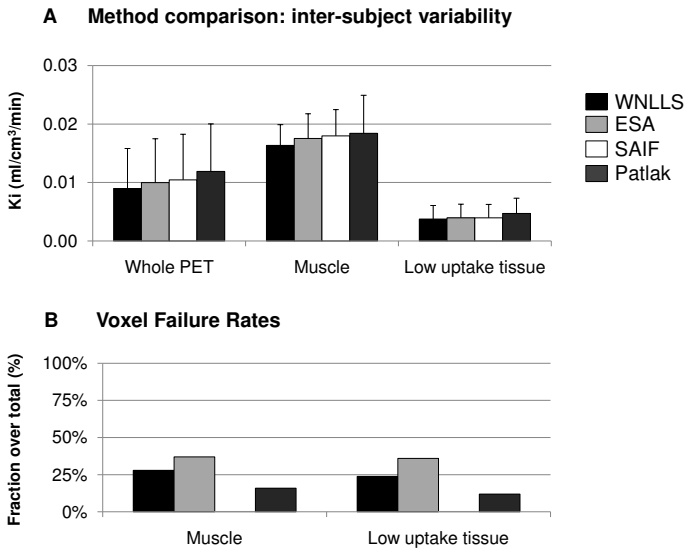

Figure 4.1.3: Method Performance. Panel A shows the mean of voxel estimates within ROIs for all the methods. Bars represents the mean+SD calculated among the subjects for the subsets of voxels in which all the methods successfully converged to physiological estimates. Methods show good agreement for all the analyzed ROIs. Panel B shows the fraction of voxel failures over the total number of voxels in each ROI. Values are the mean calculated among the subjects.

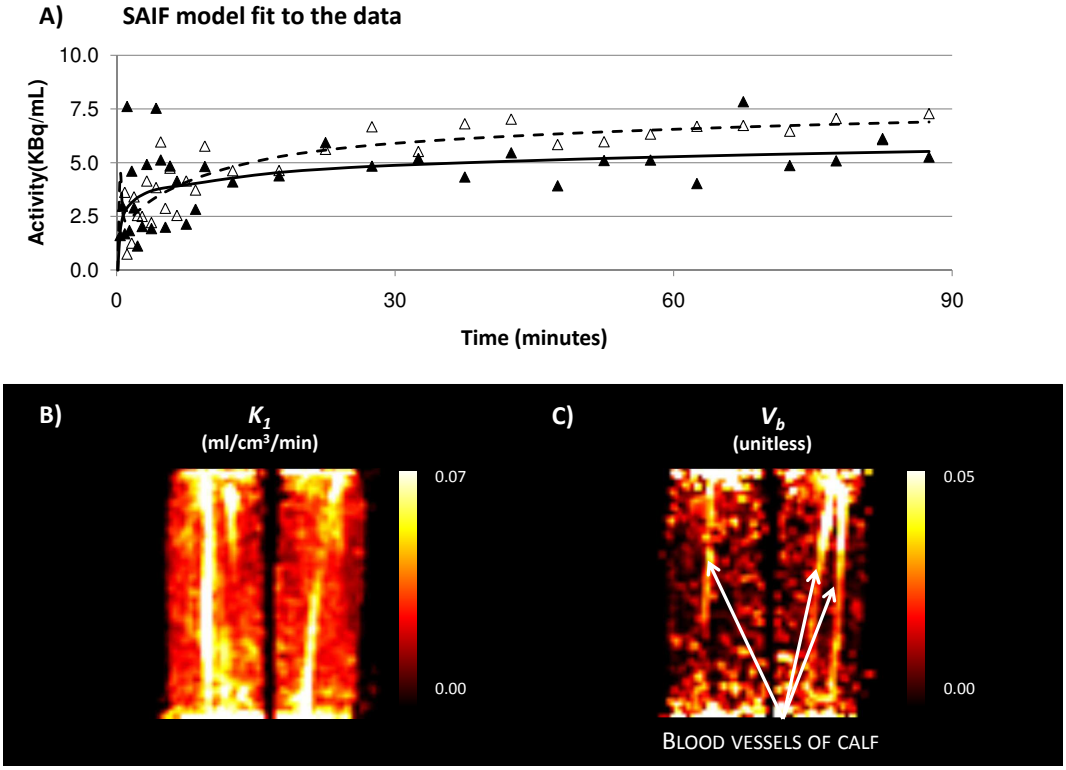

Figure 4.1.4: SAIF additional information. Panel A shows some representative voxel TACs obtained with PET-FDG applied in human skeletal leg muscle. Triangles refer to measured [^{18}F]FDG PET data (open triangles represent muscle tissues; black triangles represent low-uptake tissues). Lines refer to SAIF model prediction to the data (dashed line refers to muscle tissues; solid line refers to low-uptake tissues). Panel B and C report respectively K_1 and V_b parametric maps (coronal view), obtained with voxel-wise SAIF in one representative subject.

4.1.4 CONSIDERATIONS

SAIF demonstrated to be the best method among those tested for [^{18}F]FDG PET quantification in skeletal muscle: it provided the best K_i parametric maps with minimum failures and high anatomical resolution; respect to SUV or Patlak plot, it returned complete information about tracer kinetics, by providing for >99% of the voxels the model description of the data; as WNLLS, SAIF returned K_1 and V_b parametric maps but with a significantly lower number of outliers.
Consistently with the results obtained, it can be concluded that SAIF represents a valid and complete tool for the voxel-wise quantification of [^{18}F]FDG PET data in skeletal muscle.

4.2 [^{18}F]FLT PET STUDIES IN BREAST CANCER PATIENTS

4.2.1 INTRODUCTION

Breast cancer is the most common cancer in women worldwide and one of the principle cause of death from cancer among women globally. Despite the high incidence rates, in Western countries, 89% of women diagnosed with breast cancer are still alive 5 years after their diagnosis, which is due to detection and treatment (Parkin et al., 2008). Nevertheless approximately 30% of patients develop metastases (SEER Cancer Statistics Review, 1975-2005, `http://www.seer.cancer.gov`) and despite the advances in treatment, a too large portion of these patients will be incurable.
One important aspect to ameliorate the efficacy of the breast cancer treatment is related to the capacity of an early and object measurement of the therapy response. This possibility would allow a better calibration of the therapy to the individual, minimizing patient exposure to ineffective and potentially toxic treatment regimens. At the present time, response to chemotherapy is routinely measured with RECIST criteria (Therasse et al., 2000), which relies on changes of dimensions of target lesions. This method, however, takes time to be developed and does

not provide any indication about the kinetic changes within the tumour region. Different alternatives are already used in clinics, like histological examinations (Dowsett et al., 1999) or molecular imaging methods (Wahl et al., 1993). The PET tracer [^{18}F]FLT (fluorothymidine), for example, is a cell proliferation marker which have been successfully applied in breast cancer (Kenny et al., 2005), showing good correlation between its tissue uptake and tumour biopsies (Kenny et al., 2005). In literature, standard approaches for voxel-wise quantification of [^{18}F]FLT images are represented by the SUV and Patlak analysis, justified by the irreversibility of the [^{18}F]FLT tracer (Gray et al., 2010; Kenny et al., 2007, 2005). Both methods, however, return incomplete information about the investigated system: SUV represents a semi-quantitative approach and its results are strongly dependent from the factors used to normalized the injected dose (Kenny et al., 2005); Patlak plot, instead, provides a partial description of tracer kinetic limited to the trapping in tissues, without accounting for other important elements like tracer delivery or blood volume fraction. The latter might be a very penalizing factor, since the blood fraction in tumour regions could be much higher than 5% typically measured in the brain. It follows that blood contribute to the measured PET signal has to be taken into account in the quantification of this type of data. Some attempts have done with kinetic modelling (Gray et al., 2010; Kenny et al., 2005) but, due to the high noise of data, it was applied only at region level. Thus, a robust, reliable and informative quantification method for [^{18}F]FLT PET images is still missing.

Due to its characteristics SAIF might represent a valid solution for voxel-wise quantification of [^{18}F]FLT PET data, even though this problem is very challenging. In fact, respect to the brain, the kinetic variability of anatomical regions like breast, lung, liver, heart, results amplified, originating dynamic data with high heterogeneity. As consequence a very flexible quantification methods is needed. Moreover, in order to use the kinetic estimates to predict the response to the treatment, the quantification methods have to be first characterized by high reproducibility.

To evaluate the applicability of SAIF to [^{18}F]FLT PET in breast cancer data

we tried to consider all these issues. Thus, 1) we tested the reliability and the robustness of SAIF approach by comparing its results with those provided by the standard procedures; 2) we evaluated the reproducibility of SAIF in test/retest studies; 3) we verified the capacity of SAIF for the prediction of therapy response.

4.2.2 Dataset and methods of analysis

Our dataset was composed by 10 subjects with stage II-IV breast cancer. Briefly, all the subjects were scanned twice prior to and at 1 week after treatment, using [^{18}F]FLT tracer with an injected dose from 153 to 380 MBq. Automatic arterial blood sampling was also performed for each PET study and plasma metabolites were removed from arterial concentrations. Full details about subject inclusion criteria and the acquisition protocol are reported in (Kenny et al., 2005). As reference methods for FLT quantification we considered:

- Weighted non linear least squares (WNLLS) applied to literature reference [^{18}F]FLT compartmental model. This coincides with a 2-tissue model, composed by one equilibrating compartment and one trapping (Fig. 4.2.1) (Kenny et al., 2005). The contribute of [^{18}F]FLT metabolites during the delivery phase was neglected, as indicated by Kenny and colleagues (2005) (Kenny et al., 2005);

- Patlak graphical analysis;

- SUV defined as the measured radioactivity at 90 min, normalized by injected dose and body surface area (Kenny et al., 2005).

All the methods were applied voxel-wise and results were compared with those provided by SAIF applied at the same resolution level. K_i, K_1 and V_b were considered as parameters of interest. Mean of voxel estimates within regions and outliers were used for the computation of performance indexes. For each patient primary tumour, healthy breast and

vertebra were used as regions of interested. Liver and heart were excluded from the analysis since the different physiology of the systems might require a different description of the FLT kinetic.

Figure 4.2.1: [18F]FLT compartmental model. K_1 and k_2 refer to blood to tissue FLT transport while k_3 describes irreversible tracer uptake. The open solution of the tissue kinetic is given by $C_{tissue}(t) = \frac{K_1 k_3}{k_2+k_3}\int_0^t C_p(\tau)d\tau + \frac{K_1 k_2}{k_2+k_3}\int_0^t C_p(\tau)\cdot e^{-(k_2+k_3)(t-\tau)}d\tau$, which corresponds to the SA model composed by 1 trapping and 1 equilibrating compartment $C_{tissue}(t) = \alpha_0\int_0^t C_p(\tau)d\tau + \alpha_1\int_0^t C_p(\tau)\cdot e^{-\beta_1(t-\tau)}d\tau$.

4.2.3 SAIF PASSBAND FILTER DEFINITION

As seen from previous chapters, the most critical variable for SAIF utilization is related to the choice of the filter passband. In this applicative context the simulation-based approach previously used for passband definition was not applicable for two reasons in particular: 1) *a priori* information about the tracer kinetics in the investigated system was not available; 2) the variability of the kinetics within the analyzed field of view was too high to allow the use of a unique passband for all the analyzed ROIs. Thus a criterion to select an appropriate passband depending on the characteristics FLT uptake in the tissues was required.
Analysis of L-[1-^{11}C]Leucine SAIF passband (Section 3.2) had provided precious suggestions to address the problem, in particular for what concerns the relationship between SAIF passbands used for the region-wise and voxel-wise L-[1-^{11}C]Leucine quantification. In fact, even if the ROI and voxel analysis had shown to require different passband filters, the two entities were quite close to each other. Thus we decided to use a

hierarchical approach through which we extended the SAIF applicative setting used for the region-wise analysis to the voxel one. Since ROI analysis presents low noise and high SNR, SAIF passband is much easier to characterize, representing a good starting point for the definition of voxel SAIF setting. The criteria we implemented was as follows:

1. *For each patient, we organized the dynamic PET data in homogeneous clusters, following an anatomical segmentation. Then, for each cluster we extracted the correspondent ROI-TAC and we applied WNLLS to compartmental [^{18}F]FLT model at region level;*

2. *For each ROI, using the equivalence between [^{18}F]FLT compartmental model and spectral analysis, we computed the values of spectral components (β^*) and derived their uncertainty (SD) from the values and the precision of WNLLS regional estimates;*

3. *For each ROI we fixed the voxel-wise SAIF passband as [β^*-4SD; β^*+4SD].*

The choice of using 4 times the estimated SD was justified by the idea to be as much conservative as possible, in order to avoid the selection of a too narrow filter interval. Notably, this criterion not only provides different passbands for different ROIs but, in the case that a subject is scanned more than one time, it allows to adapt the filter passband for each PET scan.

4.2.4 RESULTS

When the methods were compared in the same subsets of voxels in which all provided reliable and physiological estimates, there was a good agreement between the results (Fig. 4.2.2A). This findings was verified for all the analyzed ROIs and all the subjects. Similarly, mean of voxel estimates within ROIs resulted well correlated each other (R^2>0.95).
The main differences between method performances were related to the voxel failures rates (VFR) (Fig. 4.2.2B): WNLLS proved to be most sensitive to noise in the data (VFR >50% in healthy breast and >20% in

tumour); Patlak and SUV, instead corresponded to the most robust methods, especially in tumours and vertebra with an average VFR of 10%. SAIF performances improved with the increasing of [^{18}F]FLT uptake: in healthy breast, where the cellular proliferation is normally low and FLT uptake limited, SAIF VFR was equal to 20%±22% (mean±SD); in tumour and vertebra, where the FLT uptake is 10 folds higher than in normal tissues, SAIF VFRs corresponded respectively to 9%±13% and 14%±14%. Visual comparison of K_i and SUV parametric maps (Fig. 4.2.3) confirmed the good agreement between the methods, providing similar information about FLT uptake in tissues. The only exception was represent by WNLLS, in which the high amount of outliers prevented the possibility to compute adequate parametric maps.

Reproducibility analysis, performed by comparing test/retest scans, demonstrated the repeatability of SUV (R^2=0.93) and Patlak K_i (R^2=0.97) parameters (Fig. 4.2.4A,B). Similarly, also SAIF and WLNNS methods reported a good repeatability of K_i test/retest estimates ($R^2\sim$0.90). K_1 and V_b estimates, instead, demonstrated weak reproducibility (K_1 $R^2<$0.65; V_b $R^2<$0.16) and higher test/retest variability respect to K_i (Figure 4.2.4C-E). These results were in agreement with previous analysis performed in literature (Kenny et al., 2007, 2005).

Since the K_i was the most reproducible parameter, it was also considered for assessment of therapy response. The main aim of this analysis was to evaluate the possibility of predicting the patient response to treatment, by comparing the variation of K_i in tumour region between pre and post treatment scans. For therapy responding subjects results showed a significant K_i decrease in tumour regions (p <0.05), while it was not found for non-responding subjects. Results were comparable for both WNLLS and SAIF (Fig. 4.2.5).

In the same way, the analysis of the distributions for voxel-wise K_i estimates showed a relevant shift toward lower values for responding subjects, while for non-responders pre-post treatment distributions remained almost comparable (Fig. 4.2.6). Analysis of K_i distributions confirmed also that the evaluation of only the mean K_i values could be misleading

for response to treatment discrimination. In fact, as indicated from Figure 4.2.5, there is a non-responding subject (*) which presents a substantial decreasing of mean K_i in the tumour region, typical of responding subjects. However, analyzing the distribution of K_i estimates (Fig. 4.2.6B) it is evident that the reduction of the mean did not coincide with a general reduction of the tracer uptake.

4.2.5 CONSIDERATIONS

Application of SAIF to [^{18}F]FLT PET data proved to be a complete and reliable tool for the voxel-wise quantification of this tracer. SAIF provided reliable K_i estimates, with high reproducibility and limited outliers. Respect to Patak, SAIF allows the computation of additional parameters of interest like V_b and K_1. However, these parameters were characterized by low reproducibility and did not show any relationship with therapy response. The results obtained with SAIF were in agreement with previous analysis presented in literature (Kenny et al., 2007, 2005) which had demonstrated that K_i was the most significant parameter for the characterization of the FLT kinetic in tumour regions, with good correspondence with histological samples (Kenny et al., 2007, 2005).
The strategy implemented for the selection of SAIF passband demonstrated to be appropriate for this dataset: by comparing the distribution of beta components derived by voxel-SAIF with the regional spectral components derived by WNLLS, we found very good correspondence between the two entities (Fig. 4.2.7). At the same time, there was also a high concentration of spectral components ($\sim$ 80% of analyzed voxels) at the border of passband intervals, suggesting that the choice could be probably too narrow. However, when we tried larger passbands (2 times wider than the original ones) the problem remained. Therefore, it is more reasonable to attribute these components to noise artefacts rather than

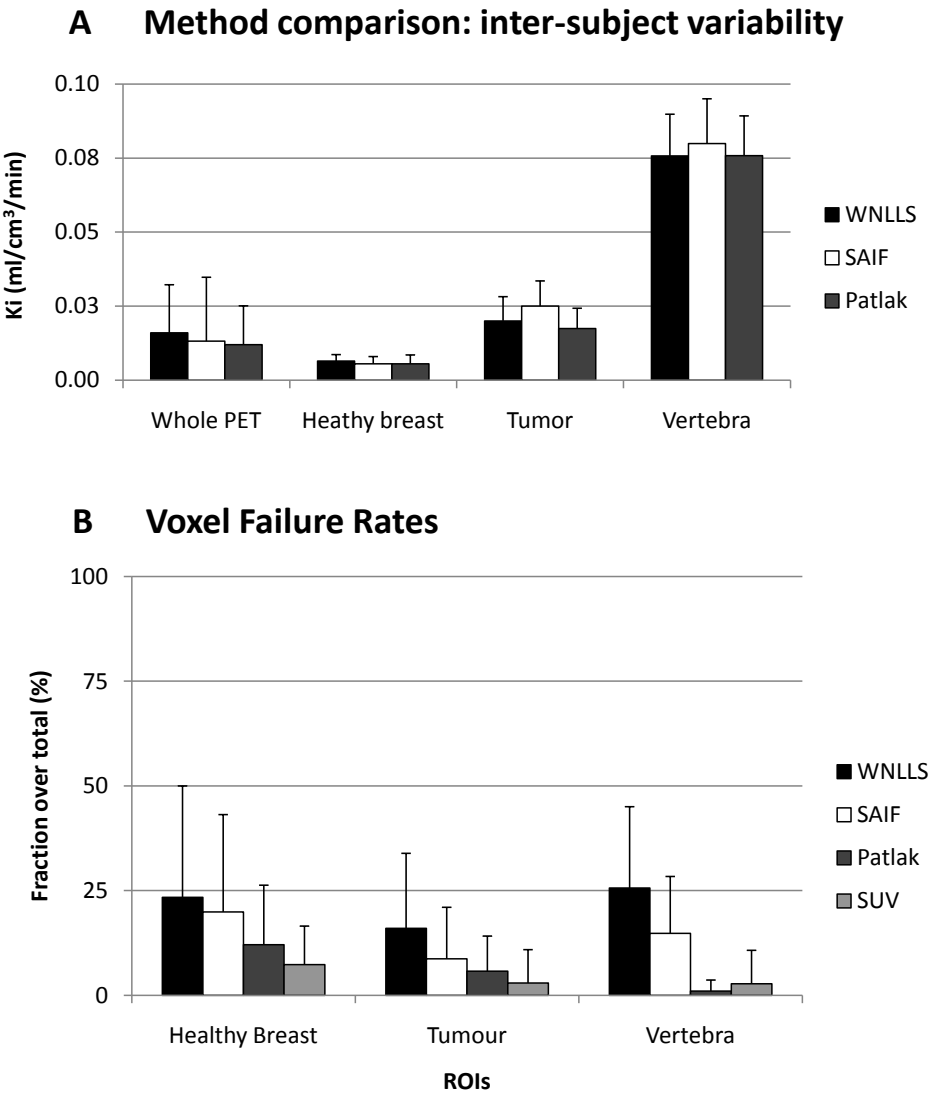

Figure 4.2.2: Method Performance. Panel A shows the mean of voxel estimates within ROIs for all the methods. Bars represents the mean+SD calculated among the subjects. Methods show good agreement for all the analyzed ROIs. Panel B shows the fraction of voxel failures over the total number of voxels in each ROI. Values are the mean+SD calculated among the subjects. WNLLS was found as the most sensitive to the noise in the data, while SUV was found to be the most robust. The limited proliferation in healthy breast tissues caused a very low signal-to-noise ratio and consequently high failure rates for all the methods.

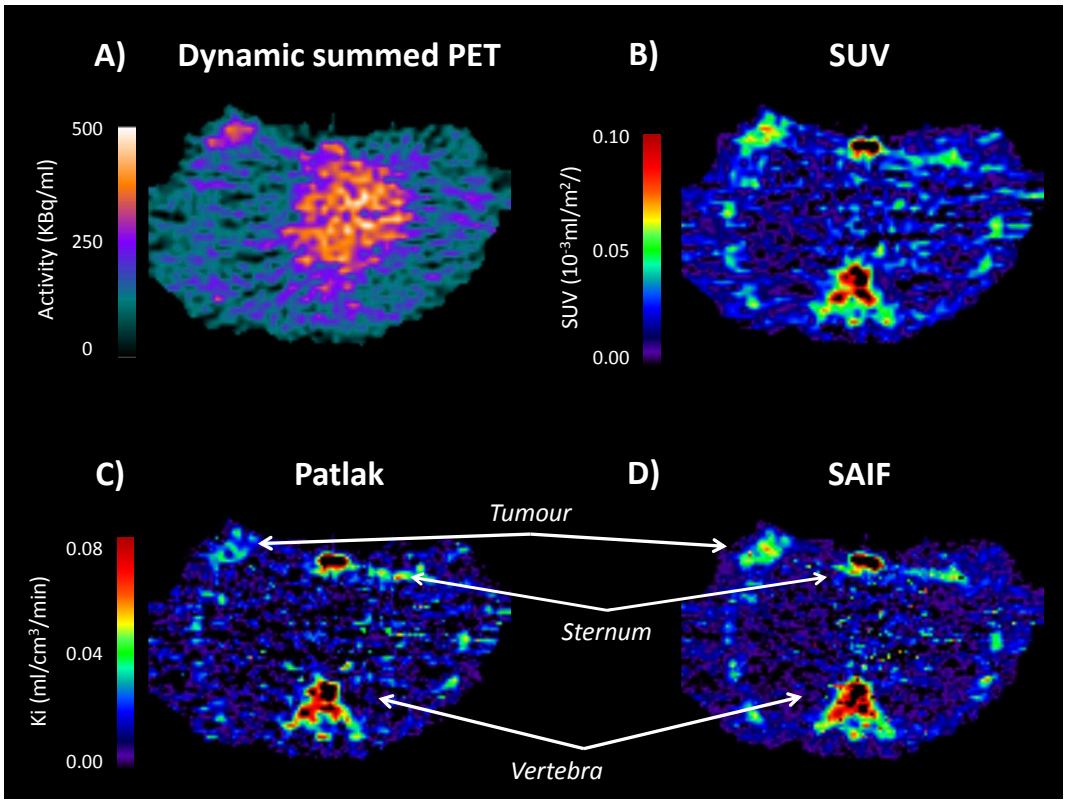

Figure 4.2.3: Parametric maps. Panel A shows the dynamic PET image summed over time. Panel B shows SUV parametric map. Panel C and D show the K_i maps obtained with Patlak and SAIF respectively. Results refer to a representative slice for the same subject. Due to high voxel failure rate (>50%), WNLLS result is not shown.

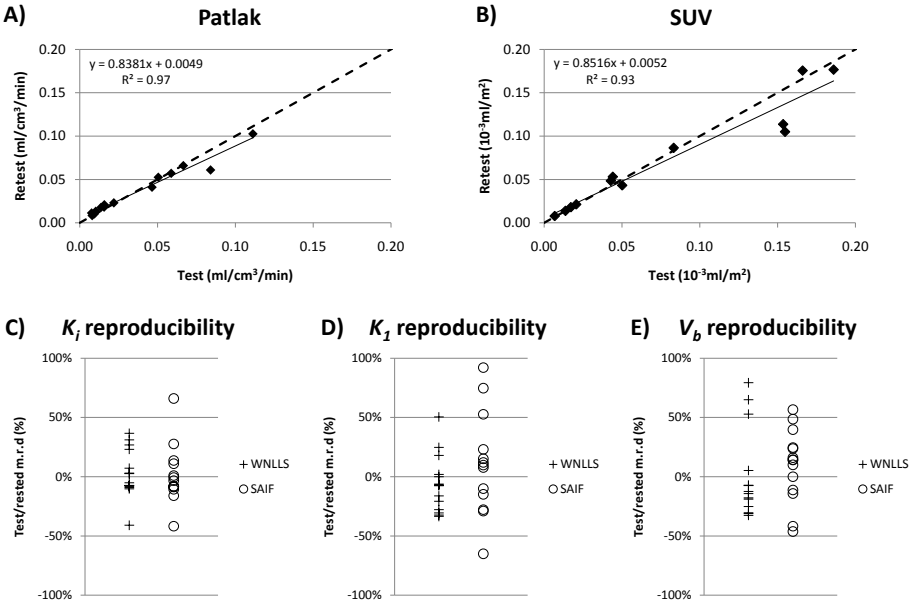

Figure 4.2.4: Reproducibility. The figure show test/retest variability of some parameters of interest. Panel A and B show respectively Patlak K_i and SUV scatter plots. Panels C, D and E show the mean relative differences between test and retest parameter estimates, obtained with WNLLS (+) and SAIF (o). Analysis was performed for K_i, K_1 and V_b considering all the subjects and all the ROIs.

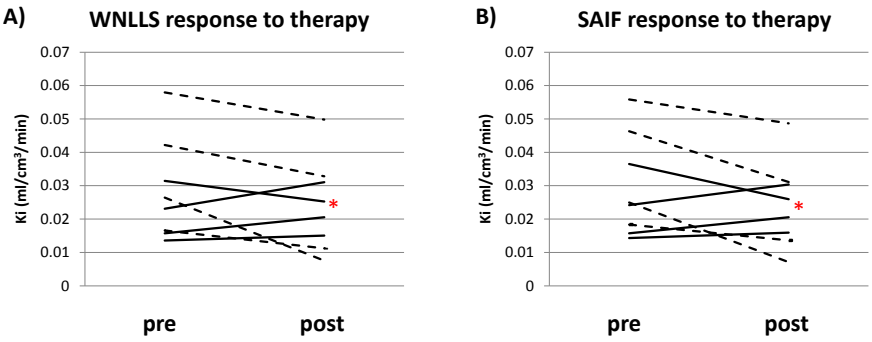

Figure 4.2.5: Response to treatment. The figure shows the change of mean voxel K_i estimates in tumour pre and post treatment for all the subjects. Dashed line=responders; Solid line=non-responders. WNLLS and SAIF results are in agreement for all the subjects. All the responders show a decrease in K_i values, while the non-responders present a general increase. The only exception is represented by *, who report a decreasing of K_i values even if non-responder.

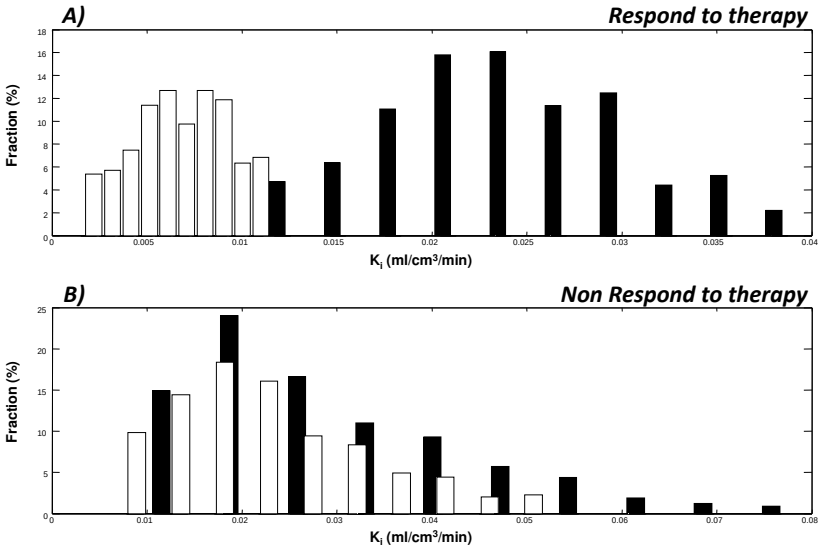

Figure 4.2.6: Response to treatment. The figure shows the change of mean voxel K_i estimates in tumour pre (black bars) and post (white bars) treatment for two representative subjects (responder and non-responder respectively). In general for the non-responder subjects the pre-post K_i distributions are overlapped, while for responders post treatment K_i distributions are shifted towards lower values. With this analysis, the subject marked with * in Fig. 4.2.5, was correctly classified as non-responder (Panel B).

real components of the system.

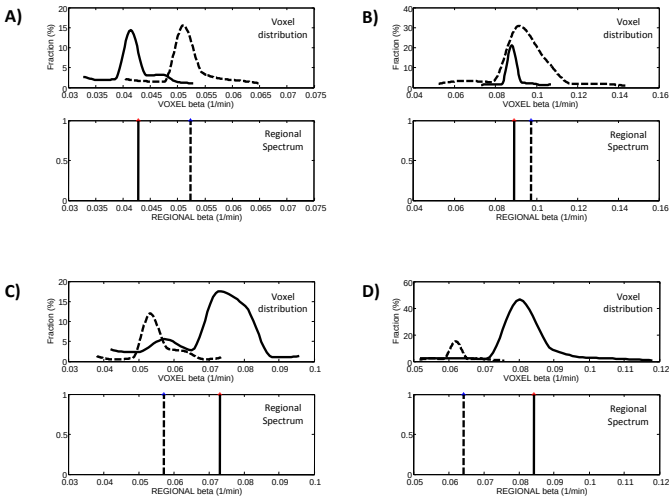

Figure 4.2.7: Distribution of beta components. The four panels of the figure show the distribution voxel-wise SAIF beta components (top of panels) with those estimated at region level with WNLLS (bottom of the panels). Four representative subjects are reported, each one considering the tumour region. For each panel, test (dashed lines) and retest (solid lines) results are visualized.

4.3 [^{18}F]FDG PET STUDIES IN ACUTE LUNG INJURY PATIENTS

4.3.1 INTRODUCTION

Acute lung injury (ALI) is a severe lung pathology that usually occurs as consequence of ordinary injuries (e.g. trauma, burns, aspiration) or an acute illness (e.g. pneumonia, acute pancreatitis) (Bernard, 2005). ALI has no specific therapy but it is common practice to treat it with mechanical ventilation in supportive care in order to reduce airway pressure while maintaining adequate oxygenation. Despite this, ALI is characterized by considerable morbidity and high mortality (Bernard, 2005; Matthay, 2008).

The main problem of ALI is that the physiological processes that are behind this evolution are not completely clear. The most shared hypothesis is that they are correlated to polymorphonuclear neutrophils (PMNs) activation, that is the mainstay of inflammation. However the uncertainty about the dynamic of the system justifies an onward interest in studying lung inflammation.
Since neutrophils activation is associated to an increase in cellular metabolism of 20-30 times respect to normal, [^{18}F]FDG -PET studies have been successfully applied for the description of the dynamic of lung inflammation (Bellani et al., 2009, 2011). With [^{18}F]FDG PET method, in particular, clinical investigation of the inflammation distribution as well as its state of activation has been possible (Bellani et al., 2009, 2011).
In disagreement with this theory, Prost and colleagues (de Prost et al., 2010) found in a sheep-model of alveolar lavage that an increase in lung water can generate an additional volume of distribution of [^{18}F]FDG . This volume is not a precursor for phosphorylation and can artifactually increase lung [^{18}F]FDG uptake independently of lung inflammation.
The source of contradiction is mainly related to lack of an exhaustive kinetic model able to fully describe the [^{18}F]FDG processes in both healthy and injured lung tissues. In these context SAIF might be a valuable investigative tool to explore the kinetics of the tracer in the tissues.

4.3.2 Dataset and methods of analysis

A dataset of [^{18}F]FDG PET images composed by 16 subjects was considered, with the aim of quantifying the dynamic of the glucose metabolism in healthy and injured lung tissues. Eleven subjects were patients affected by a severe form of ALI undergoing mechanical ventilation. Five subjects were healthy volunteer, whose studies were used as control. Normal aerated tissues (NA) and hyper-dense tissues (CO) were automatically segmented on PET images based on CT values. NA and CO were then used as regions of interest. For each subject, imaged-derived input functions were obtained from the descending aorta directly from PET im-

ages. Complete details about acquisition protocol, subject inclusion criteria and data pre-processing are reported in (Bellani et al., 2009).
SAIF quantification was applied voxel-wise and its results compared with Patlak graphical analysis, tissue-to-plasma ratio and SUV, defined as the measured radioactivity at 90 min, normalized by injected dose and body weight (Grecchi, 2011).
SAIF passband was chosen based on the distribution of spectral components obtained with ESA at region level. This resulted in a unique passband interval, for both NA and CO ROIs and all the subjects equal to [0.05,1] min^{-1}.
Comparison of K_i parametric maps as well as the mean of voxel K_i estimates with ROIs was performed to test the reliability of the SAIF methods respect to the standard procedures used in literature for [^{18}F]FDG quantification in lung, i.e. Patlak graphical analysis, SUV and tissue-to-plasma ratio (RATIO). The pseudo distribution volume V_T, defined as the distribution volume calculated on SAIF reversible components, was also computed.
Once the correctness of SAIF was assessed, the analysis of the number and the types of spectral components returned with SAIF at both ROI and voxel level was performed. The final aim was to characterize the [^{18}F]FDG kinetic model in healthy and injured lung tissues and to evaluate the presence of eventual differences between the two conditions.

4.3.3 Results

Assessment of SAIF quantification The comparison of parametric maps (Fig.4.3.1) showed that the estimates of K_i obtained with SAIF, RATIO and Patlak are overall consistent while some activated areas are lost with the SUV (in particular in the lower portion of the lung). This lack of performance is even clearer looking at Figure 4.3.2. While the SAIF exhibits the same K_i of Patlak (R^2=0.98, Fig. 4.3.2A) and similar behaviour with the RATIO method (R^2_{RATIO}=0.97, Fig. 4.3.2B), the correlation of SUV results with the SAIF ones was found not satisfactory

(R^2_{SUV}=0.67, Fig. 4.3.2C). This suggests that the more the dynamic contribution is taken into consideration, the more useful and reliable the information about the system will be. For this reason the application of the SUV could be questionable or not completely representative.

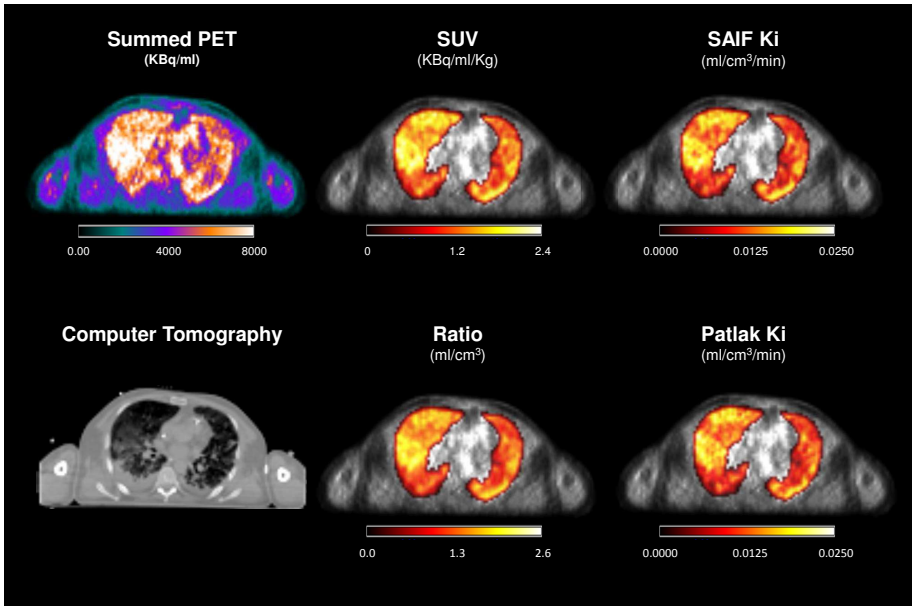

Figure 4.3.1: Quantification of [18F]FDG PET lung images. Comparison of quantification methods.

ANALYSIS OF SPECTRAL COMPONENTS Region- and voxel-wise SAIF analysis allowed to separate 3 kinetic components: the first associated with the trapping of FDG-6-P in the tissue, the second one with an intermediate reversible process, and the last one with a fast reversible plasma-tissue exchange process. In comparison with the controls, the spectra of the patients were, generally, shifted towards the right side indicating the presence of faster kinetics (Fig. 4.3.3). In particular, the degree of the shift of the faster component allowed distinguishing normal-dense, hyper-dense or edematous tissue. We found a 25-fold mean increase of K_i between controls vs. patients (0.0006 ml/cm^3/min vs. 0.015 ml/^{3}/min) and a 3-fold mean increase in V_T (0.15 ml/^{3} vs. 0.50 ml/^{3}). The higher V_T

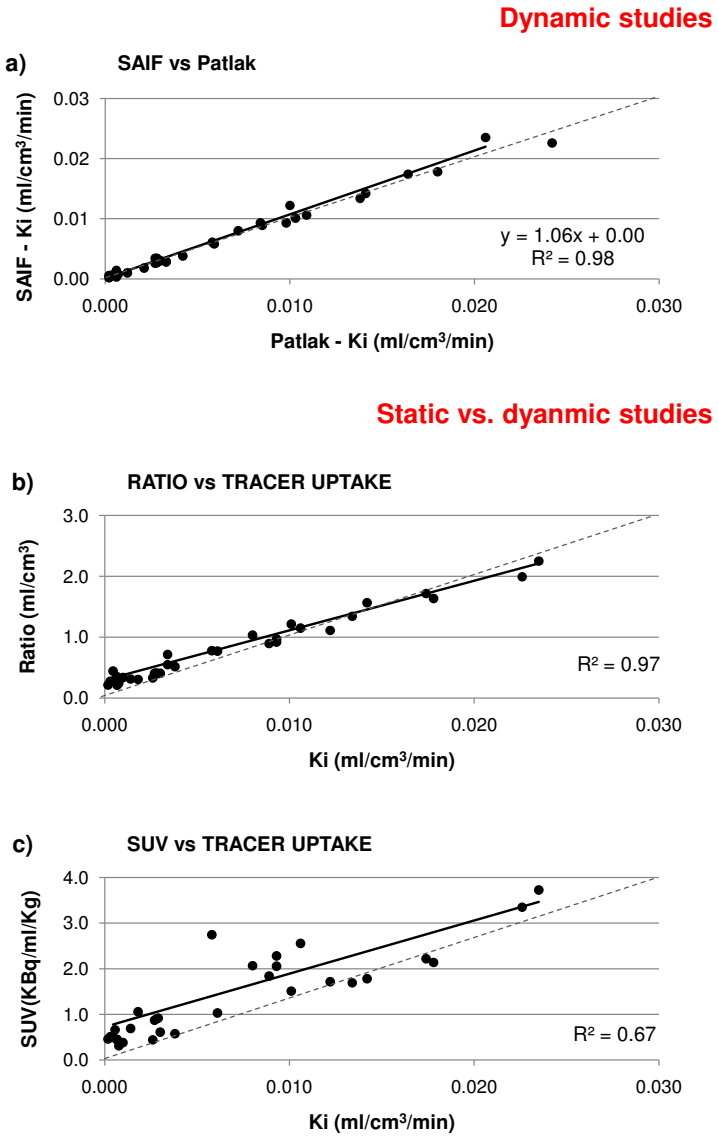

Figure 4.3.2: Method Correlations. The figure reports the scatter plots that show the correlation among the considered quantification methods. Panel A-C indicate the correlation between SAIF and Patlak, RATIO and SUV analysis respectively. Black circles indicate the mean voxel estimates within a specific region of interest (ROI). NA and CO tissues were reported for all the subjects analysed. Pearson's coefficient (R^2), slope and intercept of regression line (solid line) are also shown.

indicates an increased availability of [^{18}F]FDG in the extravascular space which is related to the pathological status of the tissue within the voxel.
Figure 4.2.4 shows the kinetic characterization provided by SAIF in one ALI patient. Respect to SUV and RATIO which only return the information about the tracer trapping, SAIF allows a more complete description of the tracer kinetic in the tissues. It can be noticed how V_b and K_1 parametric maps offer a different insight of the lung area (Fig. 4.2.4B,D) respect to K_i map, indicating that the blood volume and the tracer deliver play different roles in the inflammation process. In the same way V_1 and V_2, respectively the distribution volume of the fastest and slowest spectral reversible component, resulted different between each other and different respect to tracer uptake.
The analysis of these parameters for the entire dataset (Fig. 4.2.5) allowed a more precise classification for both control and patient groups. Notably, healthy and injured patients do not differ only in the tracer uptake but also in the [^{18}F]FDG transportation from plasma to tissue as well as in the kinetic within tissues.

4.3.4 CONSIDERATIONS

SAIF results were found well correlated with those provided by the standard quantification methods for the description [^{18}F]FDG uptake in lung, corroborating trustworthiness of subsequent SAIF estimations. This correlation was strengthened by the voxel-by-voxel analysis so that we can reasonably consider SAIF method as valid alternative for [^{18}F]FDG quantification in lung tissues.
In fact, despite the noisiness of PET data, especially at voxel level, SAIF revealed to be highly informative about tracer kinetic. The analysis of SAIF spectra showed differences between controls and patients but also between different types of pathological tissues. Looking at the results obtained at both ROI and voxel analysis we could conjecture that FDG

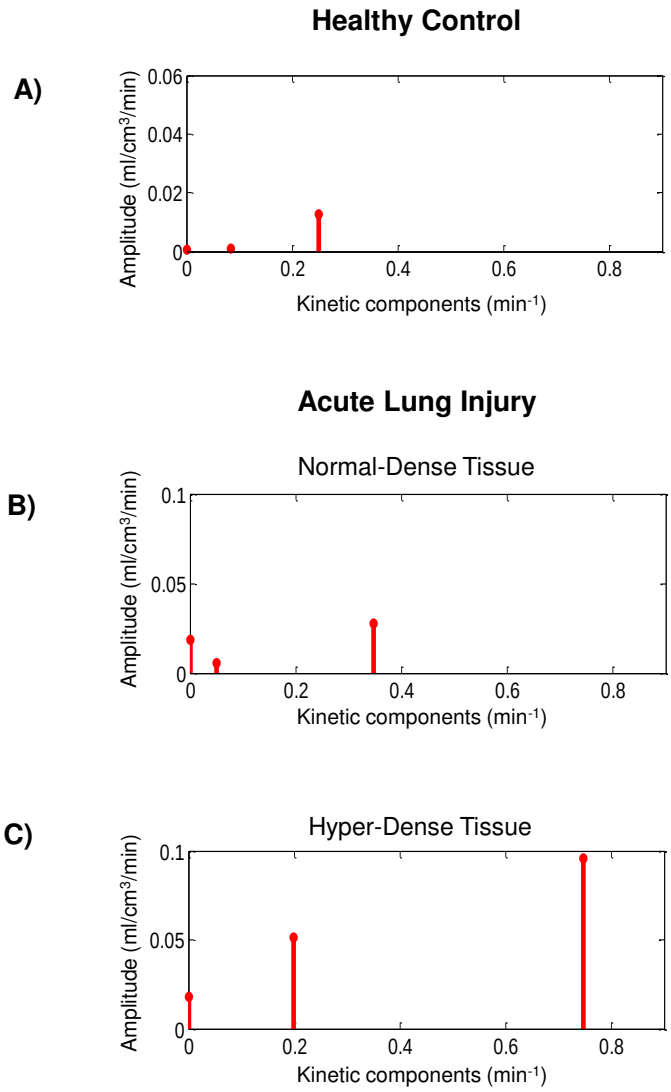

Figure 4.3.3: Spectral components of [^{18}F]FDG lung kinetics. The figure reports the spectra obtained with SAIF of two representative subjects: (A) control subject, (B) normal-dense tissue of a patient affected with acute lung injury, (C) hyper-dense or edematous tissue of the same patient affected with acute lung injury. Each line represents a component that characterizes the tracer kinetic (the x axis indicates the velocity of the kinetic component, while the y axis its amplitude). In particular the amplitude of the one in zero coincides with the uptake of the tracer in tissue. Note that the patient maintains similar K_i (i.e. amplitude of the lower component) for both normal- and hyper-dense tissues, while the latter demonstrates a significant shift of the faster components and an increase of their amplitudes (panels B and C).

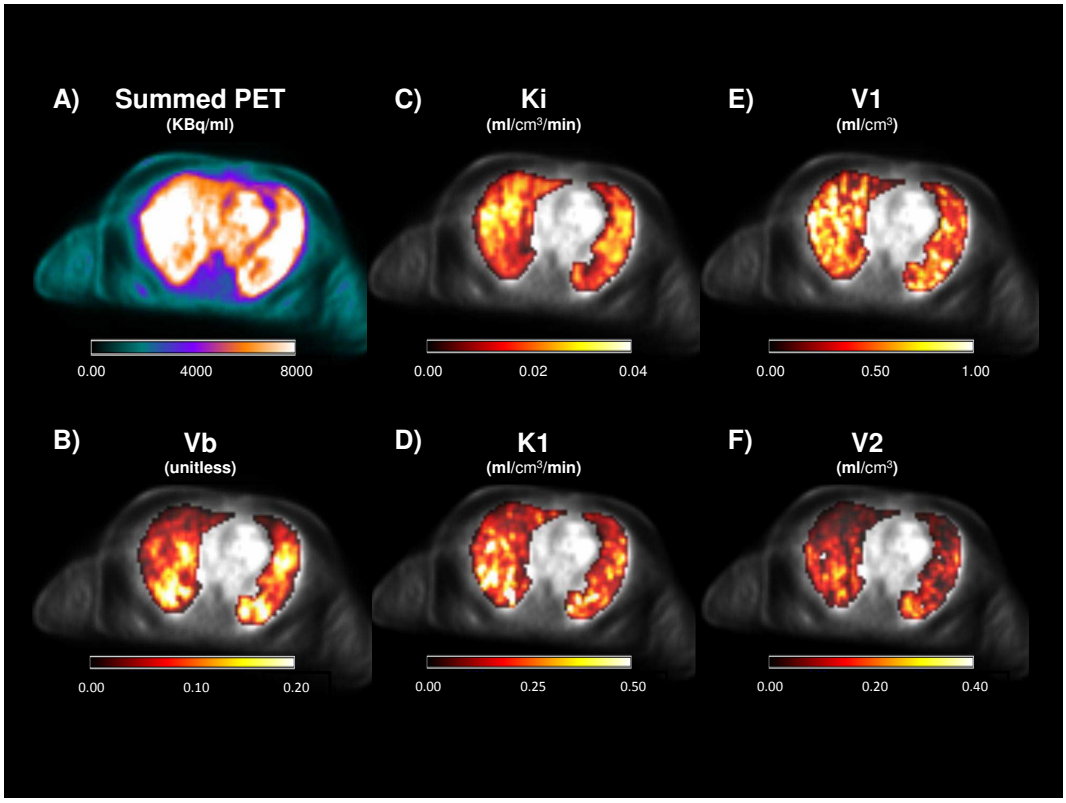

Figure 4.3.4: SAIF kinetic characterization: parametric maps. (A)Summed dynamic PET over time; (B) tracer uptake; (C) distribution volume for the fastest reversible component; (D) blood volume; (E) plasma to tissue transport; (F) distribution volume for the slowest reversible component. All the maps refer to the same ALI patient.

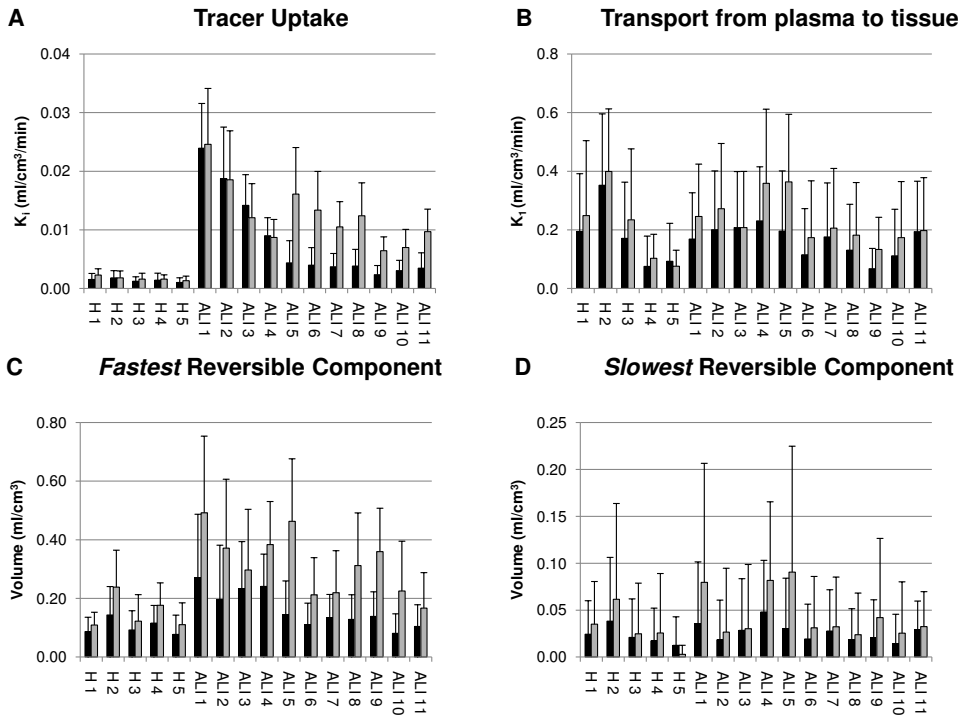

Figure 4.3.5: SAIF kinetic characterization: bar graphs. Panels A-D shoe respectively the tracer uptake, the tracer delivery and the distribution volumes for the fastest and slowest reversible component in all the analysed subjects. Black bars refer to NA region. Gray bars refer to CO region. Each bar represents the mean+SD of the voxel estimates within the ROI.

kinetic in lung tissue is described by a three compartmental model. Even though we did not have enough data to formulate the connection between compartments, the good agreement between ROI and voxel analysis suggested that this effect was improbably due to related to the tissue kinetic heterogeneity. Further analysis, using model-fixed structures needs to be done, but preliminary results would suggest to start the investigation by using a homogeneous compartmental model with two equilibrating components and one trapping. This hypothesis is in agreement with the model formulated by Schroeder and colleagues (Dittrich et al., 2012; Schroeder et al., 2008) about [^{18}F]FDG kinetic in sheep lung, even if SAIF did not provide any support to the model configuration presented by the authors.

I am not discouraged,
because every wrong attempt discarded
is another step forward.

Thomas A. Edison

5

The tissue kinetic heterogeneity problem

5.1 Tissue kinetic heterogeneity in PET domain

5.1.1 Introduction in biomedical imaging context

When a specific physiological process is investigated *in vivo* in a biological organism, it is reasonable to expect a variability of response. This variability is not *a priori* predictable as it is result of a combination of different factors, which basically depends on:

- The intrinsic biological variability of the system of interest
- The variability of the process under study within the investigated system

These two elements combined together (variability of system + variability of process) are origin of what is called as tissue kinetic heterogeneity phenomenon. Even it is not possible to quantify in advance its amount, there are some general qualitative rules with which manage it. First of all, the tissue kinetic heterogeneity has always to be contextualized respect to the particular process and system under study. Secondly, it is proved that, **the higher the complexity of a system of interest, the higher the variability observed**. This concept is based on the idea that heterogeneous systems are characterized by higher variability respect to correspondent homogeneous ones. In addition, **the variability of a specific process within a biological system is not randomly distributed but it follows the anatomical and physiological organization of the elements which composed it**. This criteria is based on the common sense that similar elements have similar behaviour.
Theoretically, it is possible to assume that the kinetic heterogeneity becomes negligible only when the system can be break up in homogeneous subsets of elements. This condition is very difficult to be reached because it requires to recognize and manage all the single constitutive elements of the system of interest (Fig. 5.1.1).
When an investigative tool like biomedical imaging is applied to explore biological tissues, the presence of kinetic heterogeneity becomes a problem especially for *in vivo* applications. Since the spatial resolution of the imaging instruments is generally several order of magnitude inferior to the cell dimensions which composed the tissues [1], it is not possible to decompose the observation into its constitutive sources. Thus it is necessary to introduce the assumption that within the minimum volume of observation provided by the instrument, the voxel, the particular investigated process can be approximate by its mean value (Fig. 5.1.2). This assumption is less restrictive that what it appears, because, following the principle that heterogeneity degree is function of the tissue spatial organi-

[1] The spatial resolution for *in vivo* biomedical imaging tools is measurable in millimetres, while cell dimensions are of micrometer order.

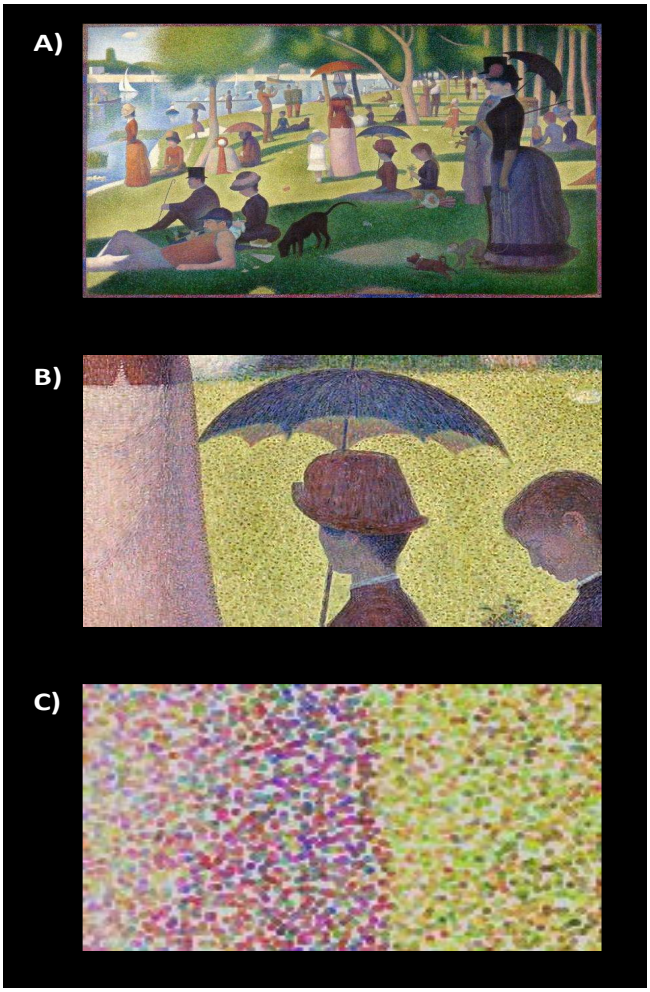

Figure 5.1.1: The Sunday Afternoon on the Island of La Grande Jatte (1886). This painting from George Seurat is one of the most famous example of pointillism, but it can be also use as a schematic representation of kinetic heterogeneity concept. From a visual analysis, it appears clear that the whole picture is the result of several graphics elements which all contribute to the final effect of the painting (Panel A). If we zoom in a particular area of the painting (panel B), the general overview is replaced by the combination of new different elements which compose the new image. If we zoom further (Panel C), we can now visualize the single constitutive elements of colour. It is following this hierarchical organization, based on the heterogeneous combination of small points of colour, that the famous painting is obtained.
In the same way, the overall response of a biological system to a specific process can be seen as the heterogeneous combination of single cell responses, that through their organization in tissues, organs and apparatus form the entire organism.

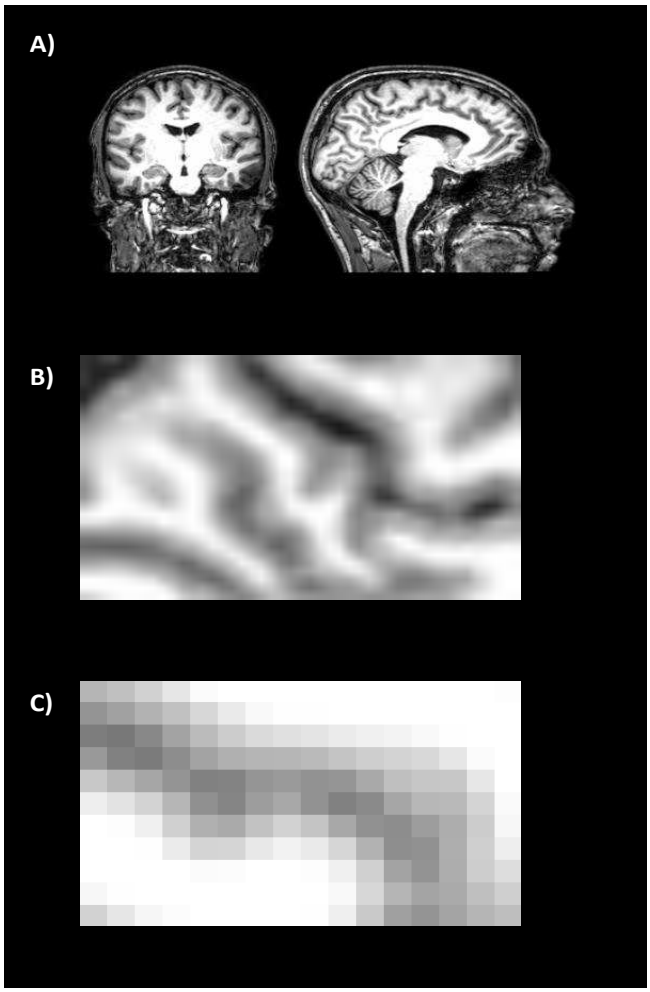

Figure 5.1.2: Heterogeneity and imaging spatial resolution. Panel A shows a T1 weighted magnetic resonance. By zooming in a particular area, like the cerebral cortex (Panel B), the whole complexity of the brain is replaced by the combination of white and gray matter. However, due to the limit of MR scanner resolution, it is not possible to zoom in until the constitutive elements of the tissues, the cells, are individually reported. The best analysis is limited at the voxel level, where each voxel represents the mean activity measured in the tissues within its volume (Panel C). This assumption can be acceptable only when the tissues within the voxel volume are homogeneous and therefore indentifiable through the mean operator. On the contrary, when the voxel contains a mixture of different tissues, like when it is located at the border between gray and white matter, the heterogeneity of the tissues must be taken into account for a correct quantification.

zation, the higher the spatial resolution is, the lower the error introduced by the average simplification. On the other hand, the presence of partial volume (PV) [2] introduced by the imaging technique becomes an additional source of error.
From these observations it follows that the impact of kinetic heterogeneity is not the same for different imaging techniques, and not only because the resolutions are different, but also because the image contents can more or less conditioned by the problem. In structural imaging, where the main goal is to characterize the anatomy of the system and not to evaluate its physiology, the presence of kinetic variability of the tissues does not represent a problem. On the other hand, in functional imaging, the presence of heterogeneity not only has to be taken into account for a correct quantification of the process, but its measures might represent an information of the system status underling the eventual presence of pathologies. Thus, case by case the presence of heterogeneity becomes an issue with different importance which has to be characterized in the particular context of interest.

5.1.2 Kinetic Heterogeneity and PET

One of the great strengths of PET is that it can be used to provide quantitative measurements of physiological and biochemical processes. The ability to determine absolute rates of a specific, well-defined biological process is essential when different individuals or groups are to be compared. In this respect, quantitative PET functional imaging methods contrast with methods in which only relative rates of a process are measured, or methods whose measures reflect an indeterminate combination of many biological processes.

[2]The partial volume effect can be defined as the loss of apparent activity in small objects or regions because of the limited resolution of the imaging system. It occurs in medical imaging and more generally in biological imaging such as PET and single-photon emission computed tomography (SPECT). If the object or region to be imaged is less than twice the full width at half maximum (FWHM) resolution in x-, y- and z-dimension of the imaging system, the resultant activity in the object or region is underestimated (Hoffman et al., 1979).

The tissue kinetic heterogeneity problem

Functional quantification with PET is generally based on kinetic modelling approaches that relate a particular biological process of interest to measurements of activity in blood and tissue following administration of a radiolobelled tracer. Kinetic models used in PET are necessarily simplified representations of tissue processes, and one of the simplifying assumptions frequently made is that tissue regions are kinetically homogeneous, i.e., rates of blood flow, delivery and efflux of tracer to/from tissue, metabolism, and incorporation into labelled products do not vary in the tissue region. In brain, for example, these assumptions are difficult to meet. As succinctly stated by Sokoloff et al (1977) (Sokoloff et al., 1977):

The mammalian brain is a complex heterogeneous organ comprising many structural and functional components with different and independently regulated levels of functional activity and energy metabolism.

At spatial resolutions approximately an order of magnitude higher than PET, such as achieved in autoradiographic studies, one clearly sees heterogeneity of rates of blood flow, glucose metabolism, and protein synthesis across the brain (Schmidt et al., 2005). Not only are rates of these processes different in gray and white matter, but within gray matter structures themselves, and to a somewhat lesser extent in white matter, rates of these processes can vary considerably. At the relatively lower spatial resolution of PET scanning, therefore, activities measured in brain can be expected to originate from kinetically heterogeneous mixtures of tissue. Application of kinetic models designed for homogeneous tissues to heterogeneous tissues has been shown to lead to errors in estimated rates of cerebral blood flow and glucose metabolism, as well as to errors in estimates of receptor binding parameters (Schmidt and Turkheimer, 2002). The same reasoning can be extended when PET is applied for the analysis of non brain tissues. As for the brain, also other apparatus are characterized by highly complex structural and physiological organizations, which required to be taken into account for a correct quantification. The problem is complicated by the fact that the degree of tissue heterogeneity,

within a given volume of observation, is not *a priori* known but it has to be determined from the data. Thus, to address the issue of kinetic heterogeneity in PET, the quantification method has not only to account for its eventual presence but, at the same time, it has also to be adaptable to different heterogeneity levels.

5.1.3 SA: A FLEXIBLE TOOL FOR KINETIC HETEROGENEITY IN PET

In Chapter 2, it has been demonstrated that SA model has the potential to be applied to homogeneous as well as heterogeneous tissues without any a priori assumptions. Moreover by evaluating the number kinetic spectra estimated, it can be defined the number of sub regions composing the system of interest and thus it is possible to identify the heterogeneity degree of the particular volume of interest. These two features of SA combined together offered the potential to exhaustively face the problem of heterogeneity in PET.
Unfortunately standard SA has shown to be sensitive to the noise in the data, characteristic that not only prevents its application at low SNR level, but which also conditions the capability of the method to correctly manage and measure the kinetic heterogeneity of the tissues. Since SAIF has demonstrated a superior robustness to the noise, especially when applied voxel-wise, it might represent a better tool for tissue kinetic heterogeneity measurement in PET. In the next sections we will test it, by evaluating SAIF performance in simulation as well as on L-[1-^{11}C]Leucine measured PET data.

5.2 MEASUREMENT OF KINETIC HETEROGENEITY IN SIMULATION

5.2.1 DATASET

In order to test the SAIF capability to correctly manage the tissue kinetic heterogeneity of the data, the same simulation procedure described in

Chapter 3 was considered.
Briefly, 75 simulated irreversible tracer kinetics were used to define noise-free tissue TACs. Both homogeneous and heterogeneous spectra were considered. For each simulated TAC, two Montecarlo simulations were performed, respectively by using region-wise and voxel-wise simulated SNR levels (for further details about the simulation procedure, please refer to section 3.1.1).
SAIF was then applied to the simulated data and, in order to test the SAIF correctness on measuring the heterogeneity degree of the data, the number of equilibrating components estimated from the data was compared with the simulation reference values. The analysis did not include doubling components, which were removed from the estimated spectrum through correction previously described in Sect. 2.3.2. The means of K_i, K_1 and V_b bias% obtained from homogeneous simulated kinetic spectra were compared with the correspondents obtained from heterogeneous ones. This analysis was performed to evaluate the impact of the kinetic variability on SAIF quantification. Since the application of ESA did not provide reliable results in both term of precision and accuracy, the analysis was limited to SAIF only.

5.2.2 Results

At ROI noise level SAIF demonstrated good precision in the measuring the simulated heterogeneity degree: homogeneous spectra were correctly detected for 72% of the simulated kinetics, while the heterogeneous ones for 89% of simulated cases. To note that, as expected from the behaviour of SA-based methods (Cunningham et al., 1998), also SAIF tended to overestimate the number of components present in the data. In simulated homogeneous conditions, 22% and 6% of the SAIF estimates respectively reported 2 and 3 equilibrating components. In simulated heterogeneous conditions, instead, 9% of estimates reported the 3 or more equilibrating components. Only 2% of the heterogeneous simulated cases were estimated as homogeneous.

This tendency of SAIF to overestimate the heterogeneity of the tissues was confirmed and even amplified at voxel noise level. Only 49% of the homogeneous simulated spectra was correctly measured, while 49% and 2% was associated with 2 and 3 equilibrating components. On the contrary 79% of the heterogeneous simulated spectra was correctly measured, while the remaining part was associated with 3 or more equilibrating components. Less than 1% of heterogeneous simulated kinetic was identified by SAIF as homogeneous. These results suggested that the SAIF measurement of kinetic heterogeneity has to be very carefully utilized, especially at low SNR levels.
In term of heterogeneity impact on SAIF estimates, very good results were found: as indicated by Table 5.2.1, parameter bias between homogeneous and heterogeneous results was comparable at both region and voxel noise level. The highest discrepancy (∼4%) was found for the quantification of K_1 at voxel level. The same parameter presented also the highest bias% for K_i and K_1 in homogeneous simulated conditions.

Table 5.2.1: Bias% variability between homogeneous and heterogeneous simulated spectra

	ROI noise error						**Voxel noise error**					
Parameters	HOM			HET			HOM			HET		
K_i	-0.70%	±	2.30%	0.20%	±	2.20%	-5.10%	±	2.20%	-3.10%	±	2.10%
K_1	1.60%	±	0.80%	1.40%	±	0.80%	5.80%	±	1.60%	1.60%	±	1.80%
V_b	0.70%	±	0.50%	0.90%	±	0.50%	-1.10%	±	1.70%	2.00%	±	1.60%

5.2.3 TOWARD A NEW METHOD FOR MEASURING TISSUE HETEROGENEITY

Results of SAIF in simulations has demonstrated that even though the quantification of the parameters of interest is not influenced by the homogeneity/heterogeneity of the tissues, the method is not reliable for measuring the degree of heterogeneity present in the data. This feature does not represent a restriction for SAIF applicability, but, since PET quantification is based homogeneous kinetic assumptions, it might be of high

interest to evaluate the distribution of kinetic heterogeneity within the volume of observation. From this analysis, a more appropriate modelling approach for the biological system under study might be developed.
In SAIF the major problem of measuring the degree of kinetic heterogeneity is related to the noise in the data whose effect is to increase the measure of heterogeneity. To limit this effect, we might think to repeat the measure more than one time. Then we could choose the heterogeneity degree of the kinetic under study as the most probable heterogeneity degree measured in all its realizations.
This approach is very simply to be realized in simulation. For each noise-free TAC, it is sufficient to compute N different noisy realizations and then to calculate its heterogeneity degree as the most probably heterogeneity estimated from all the N noisy realizations.
Respect to the single application of SAIF, the application of this multi-analysis in simulation showed an increasing of SAIF performance in correctly measuring the degree of kinetic heterogeneity (Fig. 5.2.1). In particular, increasing N (number of noise realizations per TAC) the fraction of successfully characterized cases increased, while the number of under and over estimated cases lowered. Notably, the figure shows also a large fraction of simulated cases (almost 80%) that is not associated with a homogeneous or heterogeneous spectrum. This is possible because, if the probability to be homogeneous or heterogeneous was not markedly toward a given direction, the correspondent heterogeneity degree measure was high uncertain. Thus we preferred to not associate these cases with any heterogeneity level. On the other hand, when the probability to be heterogeneous was higher than 75%, the correctness of the measure was almost absolute ($>$95%).
By considering these results, it appears evident that multi-analysis SAIF cannot measure the entire heterogeneity of the data. However, when the heterogeneity is detected we had high probability to take the correct measure, representing a good starting point for the heterogeneity analysis.

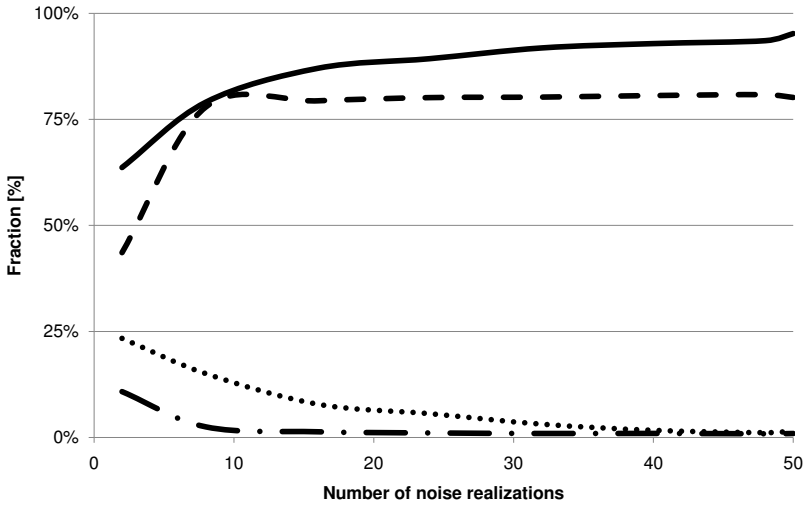

Figure 5.2.1: Performance of SAIF noise resampling method for the heterogeneity measurement. The figure shows the performance of SAIF in measuring the tissue heterogeneity as function of the number of noise realizations. Solid line represents the rate of successful classifications; dashed line represents the amount of not classified cases. Dotted and dashed/dotted lines represent respectively the over and under estimated cases. As it can be seen from the figure the rate of success increases with the number of noise realizations, while at the same time, the number of under and over estimated cases becomes negligible. On the contrary, the number of not classified voxel remained almost constant to 80%.

5.3 Measurement of kinetic heterogeneity in L-[1-^{11}C]Leucine PET data

5.3.1 Dataset

The first critic to multi analysis SAIF approach for measuring tissue heterogeneity is that in the way it is defined it results inapplicable to measured data analysis. In fact, since the true kinetics of data TACs are unknown, the noise resampling cannot be performed. Thus a different strategy for noise simulation must to be applied. As proposed by Turkheimer and colleagues in 1998 (Turkheimer et al., 1998), a valid alternative might be represented by bootstrap approach. We will test it in the L-[1-^{11}C]Leucine PET dataset.
A subset of 6 subjects belonging to the dataset presented in Chapter 3 were considered. In addition to these, 2 extra subjects investigated with low-counts doses (respectively 0.166 mCi/kg and 0.163 mCi/kg, correspondent to ∼30-40% of the normal dose) were added. A comparison between high and low dose summed dynamic PET over time is reported in Figure 5.3.1. The inclusion of these 2 new subjects aimed to evaluate the impact the injected dose on kinetic heterogeneity measurement. To note that all the subjects were acquired by the same PET scanner, ECAT HRRT, with FWHM of ∼2.6mm and voxel size of 1.21 x 1.21x 1.23 mm^3. Complete details about acquisition protocol are reported in (Bishu et al., 2008).

5.3.2 Bootstrap approach

In order to perform the analysis of kinetic heterogeneity through the multi-analysis SAIF method, we implemented the following bootstrap approach for the noise resampling:

1. Using the optimal SAIF passband for voxel-wise L-[1-^{11}C]Leucine quantification (section 3.2.3), data measured in a representative slice of an individual brain volume, masked to include brain voxels

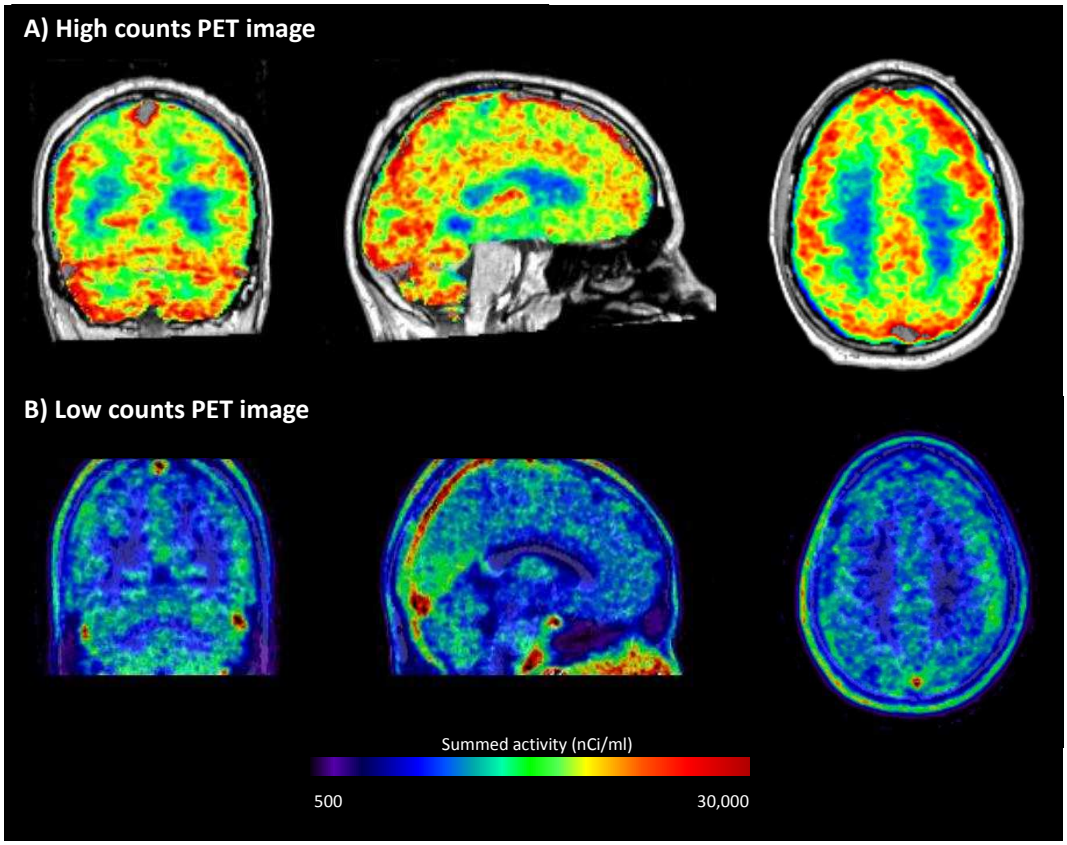

Figure 5.3.1: Comparison of High and Low count L-[1-^{11}C]Leucine Pet images. The figure shows the dynamic PET images summed over time for two subjects respectively injected with a high-activity and low-activity dose (panel A and B). Sagittal, coronal and axial views are reported.

only, were analyzed voxel-by-voxel with SAIF. The set of parameters obtained became the reference, or true, parameter values used in the simulations. Outliers and algorithm failures produced at this step were excluded from the simulation.

2. Noise-free TACs were generated by convolving the original measured arterial input function of the subject with the sum of exponential terms defined by the reference parameters from Step 1. In addition, for each voxel, normalized residuals $R_i(t)$ were computed as:

$$R_i(t) = \frac{C^*_{T,i}(t) - TAC_i(t)}{\max[C^*_{T,i}(t)]} \tag{5.1}$$

for $i = 1, 2, \cdots, K$, where $C^*_{T,i}(t)$ represents total concentration of ^{11}C in the field of view of the PET camera for the i-th voxel, $TAC_i(t)$ indicates the noise-free time-activity curve for the i-th voxel, and K is the number of voxels in the chosen slice.

3. Voxel noise was simulated using a bootstrap approach (Efron, 1979) based on random resampling of normalized residuals. With this approach, noise for the i-th voxel ($N_i(t)$) was simulated as

$$N_i(t) = R_j(t) \cdot \max[TAC_i(t)] \tag{5.2}$$

for $i = 1, 2, \cdots, K$, and $j \in 1, 2, \cdots, K$. In eq. 5.2 $R_j(t)$ represents the resampled normalized residual, originally associated with the j-th voxel, chosen by random sampling with replacement from among all brain voxels of the slice.

4. Noisy voxel TACs ($TAC^*_i(t)$) were generated by adding the simulated noise to the noise-free TACs as

$$TAC^*_i(t) = TAC_i(t) + N_i(t) \tag{5.3}$$

for $i = 1, 2, \cdots, K$

The procedure was repeated 50 times, generating 50 different bootstrap-simulated slices. Note that the bootstrap procedure used here differs from the usual approach of normalizing each residual vector by its standard deviation in order to assure that all resampled residuals have equal variance. In this approach, the required rescaling would be by the standard deviation in the target voxel residual. In the current study, however, we do not have good estimates of the variance in each individual voxel due to the inherent high noise levels of voxel data. Instead, we empirically chose normalization and rescaling to the maxima of the TACs, as these scale factors have the added advantage of preventing negative values in the simulated noisy TACs. The sample variances of the original and bootstrapped residuals were compared to check the effect of the chosen scale factors (Fig. 5.3.2).

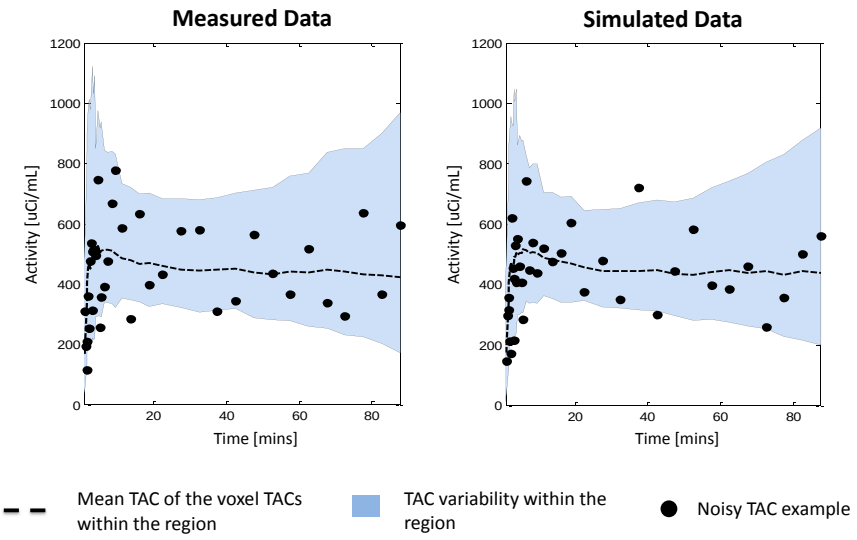

Figure 5.3.2: Residual consistency in bootstrap approach. The figure shows the kinetic variability of the voxel TACs in measured data (panel A) and in bootstrap-simulated data (panel B). The blue area represents the 90%-confidence variability interval. Black circles represent two examples of single voxel TACs in measured and simulated data. Data refer to the whole brain region of one representative subjects (high counts).

5.3.3 RESULTS

SINGLE ANALYSIS Figure 5.3.3 illustrates the spatial distribution of voxels in which two equilibrating components were detected in a representative transaxial slice at the cortical level. Results refer to single application of SAIF without any bootstrap approach. From the distribution it appears evident that the concentration of kinetically inhomogeneous voxels is randomly distributed across the entire brain slice. From a visual examination, a higher concentration of these elements seems to be located at the border between gray and white matter than within the gray or white matter itself. However, when the heterogeneity fraction was compared between gray, white and mixed gray-white matter regions no statistical difference was found in its distribution. A similar pattern was found in all subjects and brain slices, both low and high-counts subjects. By considering the results obtained in simulation, we cannot exclude that this analysis is penalized by the sensitivity of SAIF to the data noise.

BOOTSTRAP ANALYSIS Figure 5.3.4 illustrates the spatial distribution of inhomogeneous voxels obtained with bootstrap approach, by applying different threshold levels: the higher the probability threshold, the lower the amount of heterogeneity measured. This is consistent with the results obtained in simulation and in particular with the fact that more selective thresholds increase the number of not classified voxels in advantage of a superior precision of inhomogeneous voxel detection.
The analysis of heterogeneity distribution obtained with a probability threshold of 75% showed a concentration of inhomogeneous voxels 3 folds more in mixed gray/white matter areas respect to homogeneous white or gray matter. The same behaviour was found for both high and low-counts subjects.

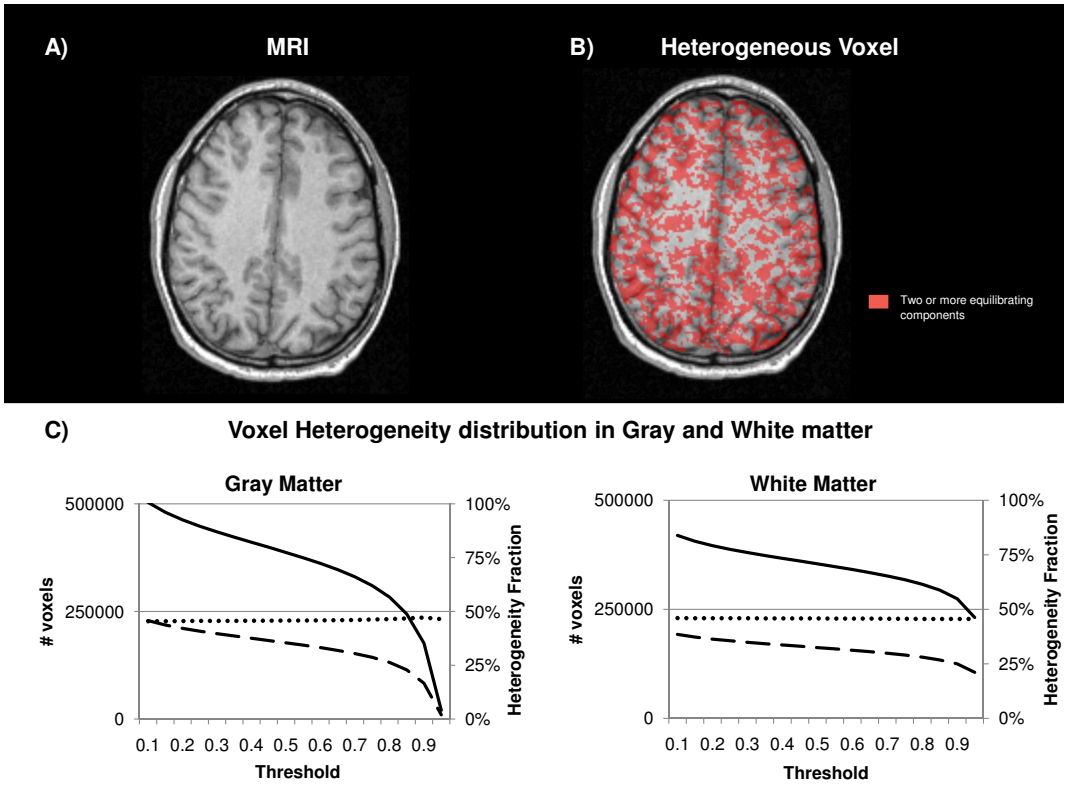

Figure 5.3.3: Heterogeneity distribution using SAIF single analysis. Panels A and B show T1-weighted MR image (left) and the same image fused with the spatial distribution of voxels estimated to be heterogeneous (right). The red area corresponds to those voxels in which voxel-SAIF detects two or more equilibrating components. Panel C show the heterogeneity fraction as function of the gray/white brain segmentation.

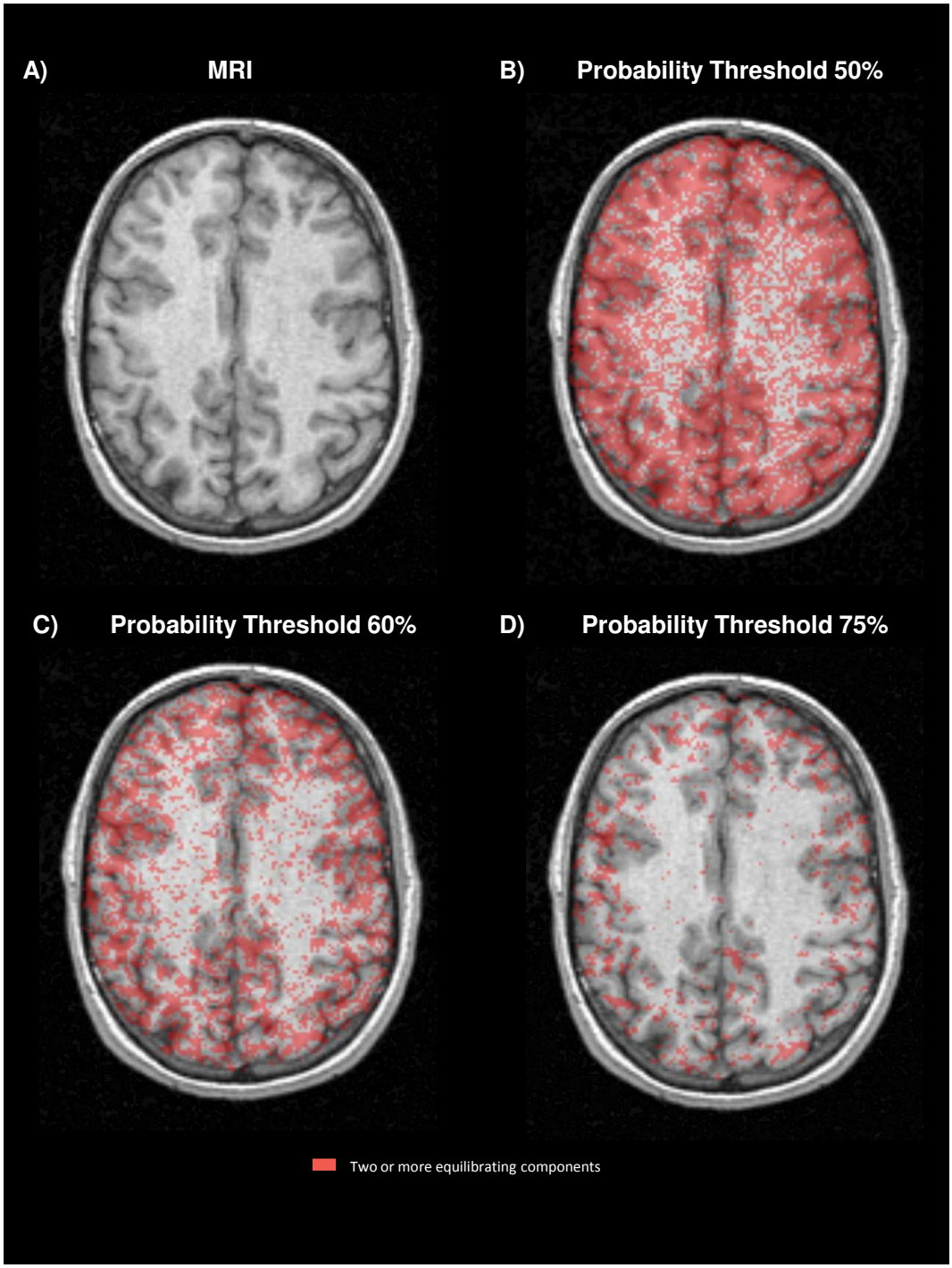

Figure 5.3.4: Heterogeneity distribution using SAIF with bootstrap approach. Distribution of voxel heterogeneity measured with SAIF by using bootstrap approach at different threshold levels.

5.3.4 THE IMPACT OF HETEROGENEITY ON QUANTIFICATION

HIGH AND LOW-COUNTS As shown from the analysis between SAIF and BFM applied voxel-wise to L-[1-^{11}C]Leucine PET image (Chapter 3), the presence of heterogeneity in the data did not significantly impact the quantification results. Even if BFM was based on homogeneous kinetic assumptions while SAIF on a more flexible model, the parameter estimates, in particular the rate of cerebral protein synthesis ($rCPS$), were highly comparable each other, both in term of spatial distribution and means within ROIs.

Extending the analysis to low-counts subjects, these results were confirmed. Figure 5.3.5 shows a comparison of the quantification results obtained with BFM and SAIF. The same good agreement detected previously was also found for low-counts data.

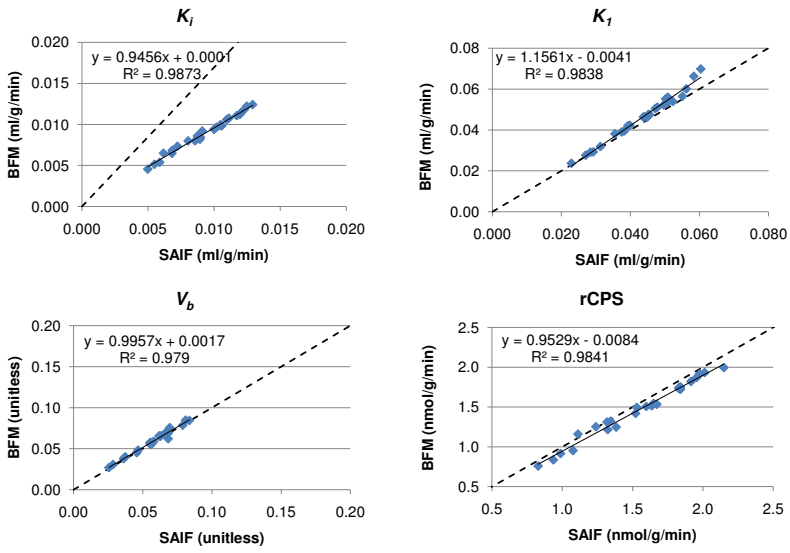

Figure 5.3.5: SAIF vs BFM scatter analysis in low-counts scenario. Scatter plots of K_i (A), K_1 (B) V_b (C) and $rCPS$ (D) are reported. The average (over all voxels in the ROI) of the parameters estimated using voxel-SAIF are plotted on the x-axis while the average (over all voxels in the ROI) of the parameters estimated using voxel-BFM are on the y-axis. All subjects and ROIs are included. In each scatter plot the slope and intercept of the fitted regression line and Pearsons R^2 are reported.

DIFFERENT RESOLUTION LEVEL The agreement found between SAIF and BFM for L-[1-^{11}C]Leucine PET data might be influenced by the image resolution. The HRRT scanner used image acquisition is a very performing PET instrument providing images with very small voxel size. Since the impact of heterogeneity grows with the increase of the volume of observation, it might be possible that voxel-wise SAIF and BFM differ at lower spatial resolution.
Thus, to mimic a lower scanner resolution, we applied a Gaussian filter with FWHM of $\sim$5mm to HRRT measured dynamic PET images (one high-counts and one low-counts image). Then, filtered data were analyzed by both methods and the results compared with those obtained from not filtered data. The same setting of previous analysis was used for both SAIF and BFM. Figure 5.3.6 summarizes the results obtained for high and low-counts cases at HRRT and artificially-lowed resolution. It appears evident an increase of the mean relative difference between methods from high to low resolution level. This behaviour was found for all the parameters in both high and low-counts studies. For $rCPS$, in particular, the increase of BFM and SAIF discrepancy was higher than 100% for both high and low-counts studies.

In term of data description and rate of outliers, both BFM and SAIF maintained a limited number of failures and outliers ($<$2%). For what concerns the WRSS and RSS, the methods provided almost the same description of the data (mean relative difference $<$3%).
Analysis of heterogeneity for low resolution data showed that for $\sim$76% of examined voxels SAIF detected the presence of 2 or more equilibrating components on the estimated spectra. This result was achieved for both high and low-counts studies. Even if SAIF tends to overestimate the heterogeneity in the tissues, compared to the high resolution data analysis it represents a statistically significant increase of the heterogeneity fraction. The same finding was also confirmed by the bootstrap approach.

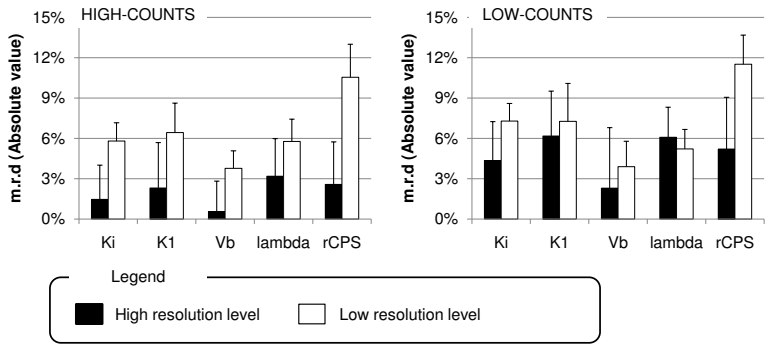

Figure 5.3.6: Mean relative difference between BFM and SAIF in high and low resolution L-[1-^{11}C]Leucine data. Bars refer to the means±SDs across ROIs, for two representative subjects (one for high-counts case, one for low-counts case). Black bars refer to HRRT normal resolution data. White bars to lowed-resolution cases. All the parameter of interest are reported.

5.4 Considerations

A correct quantification of PET data cannot be performed without accounting for the tissue kinetic heterogeneity (Schmidt and Turkheimer, 2002). In common practice, however, this aspect is generally neglected since the characteristics of the process and the system under study are such that the use of homogeneous model does not significantly impact on final result. Unfortunately this condition is not verified for all the tracers and for all the system of interest. An example is represented by L-[1-^{11}C]Leucine applied to brain.

Quantification of $rCPS$ with the L-[1-^{11}C]Leucine PET method was initially carried out at the region-of-interest (ROI) level and was based on a kinetic model that assumes the tissue ROI is kinetically homogeneous (Schmidt et al., 2005); effects of tissue heterogeneity were not taken into account and showed to lead to biased results. The problem was worsen by the fact that not all the ROIs presented the same level of heterogeneity.

To address the problem two different strategies were followed: reducing the size of tissue regions by using a voxel-by-voxel analysis, while

retaining the tissue homogeneity assumption (Tomasi et al., 2009b) and, conversely, by keeping data analysis at the ROI level, but employing a spectral analytic approach that applies to heterogeneous as well as homogeneous tissues (Veronese et al., 2010). Use of the voxel-wise analysis reduced, but did not entirely eliminate, effects of tissue heterogeneity visible in model fits of measured tissue time-activity curves (TACs) (Tomasi et al., 2009b). The spectral analytic approach, SAIF, detected heterogeneity in all ROIs examined (Veronese et al., 2010). Not all parameters were identifiable in heterogeneous tissue, however, without the use of parameter constraints. Specifically, determination of $rCPS$ requires an estimate the fraction of unlabeled leucine in the precursor pool for protein synthesis that is derived from arterial plasma, λ. In a homogeneous tissue, λ can be calculated directly from kinetic model or SAIF parameters, whereas in a heterogeneous tissue values of λ in different tissue subregions are not individually identifiable. As a first approximation we introduced the constraint that values of λ in all tissue subregions were equal, and on that basis λ and $rCPS$ in heterogeneous tissue were estimated. Under this constraint, differences between $rCPS$ estimated with SAIF at the ROI level and $rCPS$ estimated by voxel-wise analysis were small, but $rCPS$ determined with SAIF had a tendency to be somewhat lower in whole brain and cortical regions, and higher in some subcortical and white matter regions, than $rCPS$ determined by voxel-wise analysis (Veronese et al., 2010). The constraint was understood to be an imperfect approximation, as parametric maps of λ show some spatial variation (Tomasi et al., 2009b), but sensitivity analysis suggested that the use of the constraint should have only a small effect on calculated $rCPS$ (Veronese et al., 2010). Effects may vary from region to region, however, and even small biases in estimated $rCPS$ may be consequential when attempting to detect changes in $rCPS$ that are small in magnitude.

These results indicated the necessity of a L-[1-^{11}C]Leucine quantification method to deal with heterogeneity at the voxel level. The improved robustness of SAIF method to the noise in combination with the capacity of spectral analysis to take into account tissue heterogeneity have shown

voxel-wise SAIF a valid solution for this problem. As indicated by the results presented in Sect. 3.2, SAIF provided a good description of L-[1-^{11}C]Leucine data but, much more important, its results were found highly correlated with those provided by BFM. This suggested that the assumption of homogeneity used by BFM, even if not completely corrected, could be considered reliable for the voxel-wise application of this method. Moreover it has been shown how this behaviour was independent from the injected dose of tracer to the patient.
However, when the resolution of the scanner was lowered, a significant discrepancy between the two methods was found and much higher heterogeneity degree detected. Since the heterogeneity is inversely proportional to the spatial resolution of the image, in the new conditions the homogeneity assumptions used with BFM might not be completely valid.
The same reasoning can be extend to other PET tracers applied in other anatomical areas: the presence of heterogeneity is unavoidable in PET data, but the necessity to manage depends on the characteristics of the process and system under study as well as on the particular operative conditions within which the analysis is performed. Image resolution is one of the most critical elements to be taken into account. In those cases in which heterogeneity cannot be neglected, SAIF might represent a valid solution to address the problem.

Computers are useless.
They can only give you answers.

Pablo Picasso

6

Spectral Analysis for Kinetic Estimation: SAKE

6.1 The need for a SA-dedicated tool

Despite the great potential offered by standard and filtered Spectral Analysis algorithms, the use of this methodology in the PET community is limited. At the present moment, only a very selective set of published works[1] report the use of SA for PET data quantification or for the model development of new PET tracers. The main reason of this poor diffusion coincides with the lack of a unified and complete software environment for SA application. Generally SA routines are realized with in-house code whose implementation is not straightforward, requiring

[1] Based on pubmed database (`http://www.ncbi.nlm.nih.gov/pubmed`) and using as research statement: *spectral analysis* AND *PET*

advanced programming skills. This narrows the application of SA only to a restrict number of people which represent a very small subset of researchers working with PET that could take advantage from SA use.
SAKE[2] (Spectral Analysis Kinetic Estimation) is a software for the elaboration of dynamic PET data developed to overcome this limitation. SAKE has been designed as a stand-alone license-free application developed to make dynamic PET quantification via Spectral Analysis friendly usable. The program allows the analysis of a large variety of PET data at both high and low SNR and with a minimum number of assumptions. It implements the major spectral analysis algorithms (standard and not filtered (ESA), Rank-shaping (RS) and with iterative filter (SAIF)), providing an unified pipeline of analysis from pre-processing to results analysis. SAKE works through a Graphical User Interface (GUI), thus no programming knowledge is required, in order to facilitate its use also by non expert IT users.

6.2 SAKE: Features and functionalities

6.2.1 Functional Organisation

In order to create a unified system to handle the PET data quantification as well as other minor aspects of PET data analysis like the data pre-processing or the visualization of the results, SAKE has been built on three independent functional blocks. To increase the flexibility of the program and to maximize its utilization, each block can be used independently or in combination with the others. Block functionalities are described as follows.
The *pre-processing* block operates an elaboration of the raw outcomes from a PET exam preliminarily to the quantification. With this block the user can manage the analysis of arterial input function data as well as perform the correction for the measurement delay. The latter is necessary to account for the time shift between the arterial concentration at the target

[2]The last released version of SAKE is available at `http://bio.dei.unipd.it/sake`

tissue of interest and the one actually measured at the blood sampling location (generally the arm of the patient). With the pre-processing block it is also possible to apply a model-description of the arterial data based on the sum of 3 exponentials. This operation might be useful to reduce the noise in the blood measures. In addition to the analysis of arterial data, the pre-processing block offers the possibility to apply simple operations to the raw PET measurements for both region- and voxel-wise data. In particular for the PET images it is possible to compute the summed dynamic PET image over the time of experiment weighted for the length of PET time frames. The pre-processing block allows the radioactive decay correction at each level of analysis (arterial, region-based and voxel-based data) for the principal types of PET isotopes: ^{18}F, ^{11}C and ^{15}O.
The SAKE *quantification* block is the part of the software dedicated to the quantification of PET data. It implements three different types of spectral analysis methods: ESA, RS and SAIF (as defined in Chapter 2). Each method is available for the quantification of both region and voxel data. The algorithm settings as, for instance, the grid of beta components (ESA, RS and SAIF) or the filter definition (RS and SAIF) have to be defined by the user previously the application of each method.
The *analysis of results* block allows the user to manage the results obtained through the previous quantification block. For the region-based analysis SAKE offers the possibility to visualize the kinetic spectrum, the model fit of the data as well as the parameter estimates for each of the analyzed TACs. For the voxel-based analysis, instead, SAKE visualizes the parametric maps of the parameters of interest, i.e. V_T for ESA and RS, K_i, K_1, V_b and the number of components for ESA and SAIF. For each parametric map axial, coronal and sagittal views are available. The results are corrected for failures (voxels in which the algorithm failed, providing negative values or infinite precisions for the estimated parameters of interest). On the other hand, the results are not corrected for the presence of outliers (values too high to be physiological, or unreliable estimates) and this type of post-processing can be set by the user in a second time. In addition, for a selected point of the map, it is possible to compare the

SA model fit with the raw voxel TAC as well as to have the information of the parameter estimates for the particular selected voxel.

6.2.2 Data Flow

Even if the functional blocks have been built to work independently from each other, the program has been designed to manage the whole process of PET data analysis (Fig. 6.2.1). This can be realized by a sequential use of the SAKE functionalities, using the SAKE blocks one after the other. A typical PET analysis starts with the pre-processing of the raw outcomes of the study. Once the consistency of the data is verified and all the required corrections are performed, the quantification can be applied. After that, the analysis of results can be used for the extraction of the features of interest. In order to optimize this sequential use of the program, each block has been designed to provide output with the same format of the input required by the following block. Moreover, SAKE keeps in memory all the analyzed data in such way that the user can move block by block without loading each time the data input. However for those users who are interested only in spectral analysis quantification it is possible to deal directly with the quantification block. In fact the user has complete access to the output, allowing the analysis of results also with external applications. Input/output data are managed by using principally text files (*.txt*) for arterial and region of interest data and NifTi (*.nii*) or Analyze format (*.hdr/.img*) for the images. This choice was made to maximize the flexibility of the program. In addition, after each quantification, SAKE creates a own proprietary file (*.sake*) which incorporates all the information of the analysis performed. The aim of this file is to simplify the re-loading of a specific analysis, avoiding to recall each one of the analyzed files. A Matlab® file (The Mathworks Inc., Natick, MA, USA) for expert users is also stored in the working directory as well as a log file containing the setting information. The first one keeps track of all the internal results performed with spectral analysis methodologies. The latter, instead, gives quick information about the setting used for the anal-

ysis. Moreover SAKE offers the possibility to create simple .pdf report of the results obtained. All the details about the standard required by SAKE for the input data files are reported in the user manual (available in `http://bio.dei.unipd.it/sake/manual.pdf`).

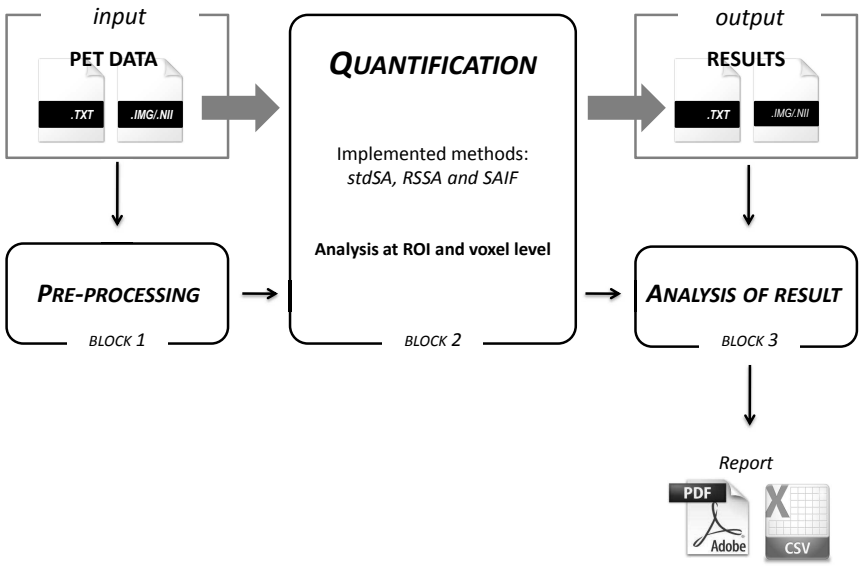

Figure 6.2.1: SAKE functional blocks. Blocks 1-3 represent respectively the pre-processing, the quantification and the analysis of results blocks. The management of input/output data can be performed through the SAKE program (black arrows) or directly by the user (grey arrows). SAKE works on *.txt* files while it supports NifTi and Analyze format for images. Results can be exported in *.csv* and *.pdf* files.

6.2.3 Software requirements

SAKE is an open-source software released under the Attribution Non-Commercial ShareAlike 3.0 Unported license (CC BY-NC-SA 3.0). The software has been developed in C++ code for Window (Windows 7/ Vista/ Microsoft Windows XP Service Pack 2+). Versions for both 32 bit and 64 bit environment are available.
SAKE works through a graphical user interface (GUI) which allows both

expert and not expert IT users to exploit spectral analysis potentials in the easiest and most intuitive possible way. Figure 6.2.2 reports some representative examples of GUI screen-shots, going through a possible pipeline of quantification process, showing SAKE's home (6.2.2A), the data pre-processing interface (6.2.2B) and the analysis of results at region (6.2.2C) and voxel level (6.2.2D).
All the pre-processing and the spectral analysis algorithms were originally developed in Matlab and then incorporated in the program through the Matlab deployment package. This choice does not impact on the way in which the user interacts with SAKE and does not obligate the user to deal with Matlab environment neither. The only requirement is represented by the installation of Matlab Compiler Runtime, which is license-free and directly downloadable with the SAKE installer.
For what concerns the system requirements, SAKE needs for $\sim$ 1 GB of disk usage and 512 MB of RAM. The resolution of the screen has to be greater than 1024x768 pixel.

6.3 Algorithm implementation

All the decisions taken to incorporate ESA, RS and SAIF in SAKE aim to maximize the adaptability of the program to the largest set of PET data. Following this principle the program leaves to the user the complete possibility to independently define the setting of each algorithm usable in SAKE. This could be critical, especially because an inappropriate choice of the setting can lead to uncorrected and heavily biased results (Turkheimer et al., 2003; Veronese et al., 2010).
The unique limitation imposed by the program is represented by the type of distribution with which the grid of components are defined. In particular SAKE realizes a logarithmic distribution for β_j as defined by DiStefano (1981) (DiStefano, 1981) and Turkheimer (1994) (Turkheimer et al., 1994).
SA and SAIF methods are based on weighted nonnegative least square

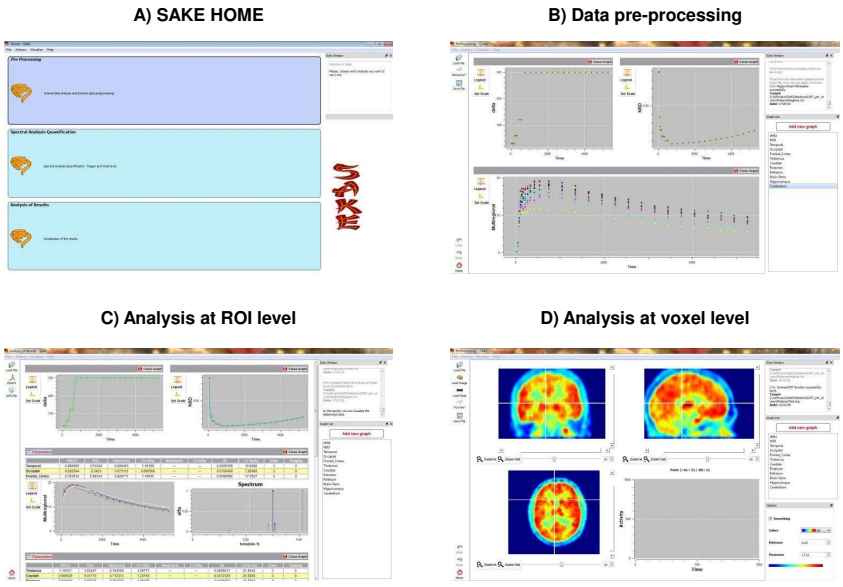

Figure 6.2.2: SAKE Graphic User Interface. The figure shows some representative examples of SAKE GUI screen-shots considering a whole pipeline of analysis: A) SAKE's home; B) data pre-processing interface; C) analysis of results at region level; D) analysis of results at voxel level.

estimator. RS, instead, for its definition requires linear estimator without the nonnegative constraint. Weights are computed inversely proportionally to the noise standard deviation of the measures. Precision of the parameters is derived as the inverse of the Fisher information matrix.

6.4 PERFORMANCES IN MEASURED DATA ANALYSIS

In order to provide some examples of SAKE functionalities, the program was tested on two different measured datasets ([^{11}C]DPN and [^{18}F]FDG , tracers with reversible and irreversible kinetics respectively). All the details about image acquisition and scanning procedure are reported in (Hammers et al., 2007) and (Bertoldo et al., 1998). Being parametric imaging of particular interest in clinics, application of SAKE was performed only at the voxel-level. In the following subsections SAKE results will be compared with the ones obtained with Patlak and Logan graphical analysis, being them the methods applied on these data in previous works.

6.4.1 APPLICATION TO REVERSIBLE TRACER: [^{11}C]DPN

Figure 6.4.1 shows the distribution volume parametric maps obtained with Logan graphical analysis (panel A), ESA (panel B) and RS (panel C) for a representative subject in a [^{11}C]DPN. The maps are reported after correction for failures and for outliers, corresponding to the voxels with non-physiological and unreliable V_T estimates. This additional correction was done a posteriori, exporting the results file provided by SAKE in Matlab. The results are in agreement with the expectation. ESA showed poor performances when applied at the voxel level, due to its high sensitiveness to noise in the data. This resulted in very noisy parametric maps (Fig. 6.4.1B) and in a 18% of outliers voxels. Both Logan and RS reported a percentage of failures and outliers $<$0.01%, being RS optimized to work with reversible tracers at high noise level. ESA and RS results were in agreement with the ones obtained in previous works (Hammers et al., 2007). As expected due to the low SNR of the data, Lo-

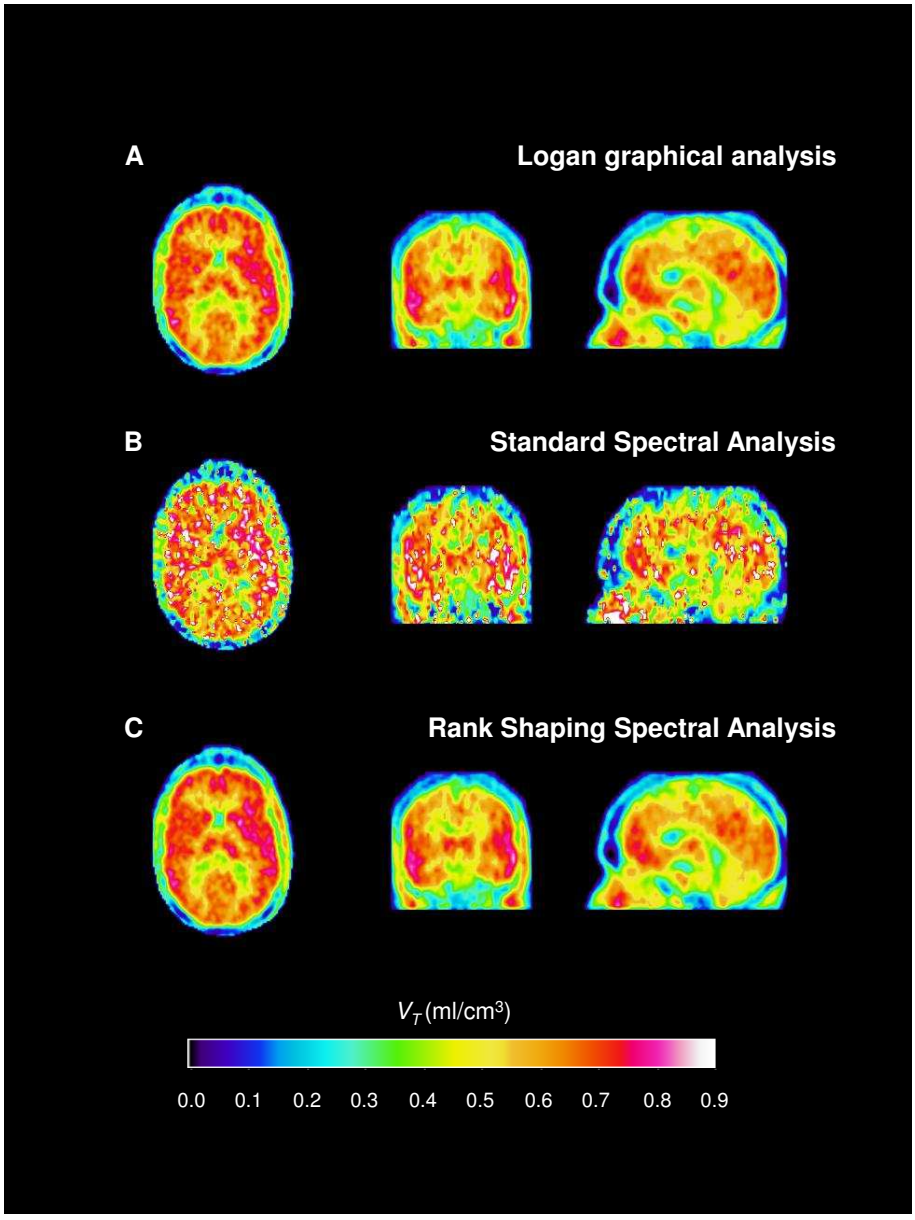

Figure 6.4.1: Parametric imaging in a [^{11}C]DPN brain study. Panels A-C report respectively the V_T parametric maps obtained with Logan graphical analysis, ESA and RS. All the three views, axial coronal and sagittal are reported. Results refer to a representative subject.

gan plot underestimated the V_T compared to spectral-based algorithms (Slifstein and Laruelle, 2001). Moreover Logan analysis is quite dependent on the choice of the time after which the plot shows a linear trend and this was a problem in particular for [^{11}C]DPN, a regionally heterogeneous tracer (Hammers et al., 2007). Thus in general [^{11}C]DPN is not quantified using Logan analysis, but alternative bias-free solutions to the Logan plot can be employed (like Likelihood Estimation in Graphical Analysis, (Ogden, 2003)). In terms of computational time, ESA required almost 15 minutes for a single subject voxel-wise analysis, while RS took around 120 minutes for the same subject. To note that Logan analysis (implemented with an in-house code) took <5 minutes for the whole quantification. These times were reported with desktop computer with 4 Gb of RAM and a 3.17 GHz processor.

6.4.2 Application to irreversible tracers: [^{18}F]FDG

Figure 6.4.2 reports the K_i parametric maps obtained with Patlak graphical analysis (panel A), ESA (panel B) and SAIF (panel C) for a representative subject in a [^{18}F]FDG study. The maps are reported after correction for failures and outliers (as defined above). Also in this case, the results were in agreement with the expectation, with ESA showing the worst performance both in terms of failures ($\sim$ 10%) and agreement with Patlak plot results (R^2 = 0.8, mean relative differences: -15%$\pm$15%). On the other hand, SAIF estimates showed excellent agreement with Patlak results (R^2 = 0.99, mean relative differences: 7%$\pm$12%), proving to be robust also when applied to low SNR data (fraction of failures and outliers $\sim$ 11%). In terms of computational time ESA and SAIF performed comparably with an average time <20 minutes for a whole subject voxel-wise analysis. To note that, as for Logan analysis, Patlak quantification (implemented with an in-house code) took <5 minutes. These performances were reported with desktop computer with 4 Gb of RAM and a 3.17 GHz processor.

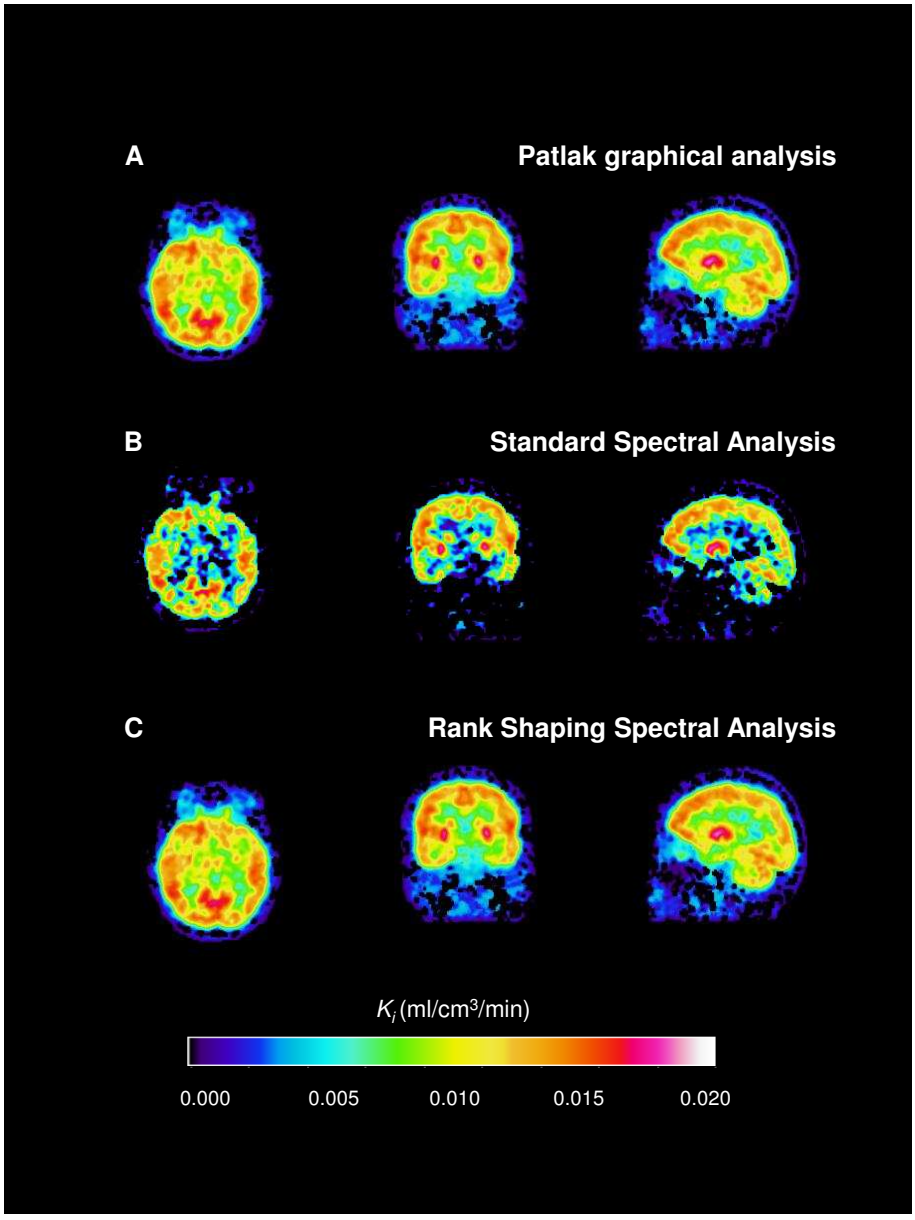

Figure 6.4.2: Parametric imaging in a [^{18}F]FDG brain study. Panels A-C report respectively the K_i parametric maps obtained with Patlak graphical analysis, ESA and SAIF. All the three views, axial coronal and sagittal are reported. Results refer to a representative subject.

6.5 CONSIDERATIONS

Preliminary studies (Rizzo et al., 2012c; Veronese et al., 2012a) have demonstrated the reliability of SAKE as a tool for PET quantification via Spectral Analysis. SAKE has been specifically designed with the possibility to manage the whole pipeline of analysis (from raw data preprocessing to analysis results) without the use of any external applications. Unlike other programs for PET image processing, only a very basic set of functionalities is available. But, exactly for this reason, it has been possible to define a simple and intuitive tool focused to SA applications. The main core of SAKE program is represented by SA algorithms. In order to maximize the utilization of the program to different tracers and different biological systems, SAKE leaves to the user the freedom to set independently the algorithm variables. Thus for a correct use of SAKE a basic knowledge in kinetic modelling is required. Theoretically this should not be a limitation because SAKE has been thought for the PET community as a simple and intuitive tool for spectral analysis utilization.

If we had had more time for discussion we should probably have made a great many more mistakes.

Leon Trotsky

7 Overall discussion

7.1 Advantages of SA quantification methods

7.1.1 Flexibility of the model

The main strength of SA methods is related to the flexibility of the model implemented: due to its additive formulation, SA can be applied to reversible/irreversible kinetics, single o multi-compartment models and homogeneous as well as heterogeneous systems. This characteristic makes SA adaptable to different tracers and different physiological systems without any *a priori* assumptions concerning the number and the types of compartments necessary to describe the data.
SA, however, is characterized by well-defined applicability limits which derive from the non negativity constraint used to estimate the coefficients of its basis function (Sect. 2.1.1). The main drawback of this assumption is the restriction of SA applicability only to models with a unique input

function and without cycling connection (Schmidt, 1999). Most of the models used in PET met these conditions, but unfortunately reference regions models do not belong to this category[1].
Respect to standard SA, filtered versions could be characterized by different (generally smaller) application domains. SAIF and RS, for example, respectively require the irreversibility/reversibility of the tracer kinetic. Thus, this characteristic has to be verified previously to their applications. RS, however, can be applied to reference region models but only for the distribution volume estimation. The larger applicability domain of RS is offered at the cost of a decreased number of outcomes.
Within its applicability domain, SA has demonstrated not only to be adaptable to different conditions, but also to be easily modifiable to the specific characteristics of the system under study. This is the case of what was done with [^{11}C]SCH442416 tracer (Sect. 3.3) or with the Double Input SA (DI) (Tomasi et al., 2012). In the first case the structure of SA model was changed to account for a vascular trapping; in the second case, instead, SA model incorporated the presence of metabolites within the tissues under study. In both situations the structure of SA model was enriched with additional components whose amplitudes were then estimated from the data as the other ones already present in the its functional basis.

7.1.2 High informative description

SA represents a complete quantification tool for dynamic PET quantification. Where it can be applied, SA allows to describe the entire kinetic of the tracer in the tissues, providing not only the information about the macro-parameters of interest, like trapping or tracer transport, but defining a full-characterization of the kinetic components necessary to describe the data. As for compartmental modelling, SA returns the model description of the entire data time-course, but without requiring a fixed

[1] It is important to stress the concepet that SA cannot be applied to reference regions, not because reference region models are inadequate to be described by sum of convolution terms, rather instead because it is not possible to estimate their kinetic spectra through the non negativity constraints

model structure. SA can be also used for model development, for the identification of the number and type of compartments necessary to describe a given system.
Not all the SA-based methods maintain the SA features, but SAIF has be designed to do it. SAIF can be applied both region or voxel-wise, providing parametric maps for K_i, K_1 and V_b estimates. In presence of heterogeneous systems, SAIF can be used for spatial analysis of heterogeneity distributions. Since SAIF is a filtered version of SA, its use is not advisable for the model development because the filtering procedure might lead to an incorrect results. SAIF is mainly focused for dynamic PET quantification, and as shown from Chapters 3 and 4, it represents a valid methodological alternative to be used for irreversible tracer analysis.

7.2 Criticality of SA-based quantification methods

7.2.1 Necessity of Arterial Input Function

For its definition SA requires the assessment of the tracer concentration over time in arterial plasma. This information is obtained by arterial blood sampling, which could be performed manually by an operator or automatically by an appropriate device. Arterial line, however, represents a risk of the patient (like infections, strokes, ...) and a risk for the personnel (risk of handling the blood of the patient, exposure to radiation,...), characteristics that limit its use in common practice. The last statistics offered by a large centre like the National Institutes of Health explained that among last 3000 PET studies performed with arterial line, only 2 patients reports problems for blood sampling and all these cases were easily solved by the personnel (Zanotti-Fregonara et al., 2011b). This might results in contradiction with the previous statement, but arterial sampling is not a complete safe practice, and for some particular categories of patients, like infants or patients with severe cardiovascular diseases for example, the presence of arterial sampling prevents use of PET. Arterial input function (AIF) is not a requirement only for SA: this fea-

ture is shared with the most used quantification methods, like graphical analysis and compartmental modelling. A valid alternative to arterial sampling could be the use of reference region, but not all the PET studies can applied these models because reference region is not always present (like in the case of L-[1-^{11}C]Leucine) and even if it is, it may be not corrupted.
With the development of non invasive input function techniques, some alternatives to the arterial sampling are present. Briefly non invasive input function techniques can be organized in 4 categories (Fig. 7.2.1):

1. **Image-derived input function (IDIF)**: the whole blood TAC is extracted directly from the PET images, by using the measured radioactivity within largest blood vessels in the field of view (Chen et al., 1998; Zanotti-Fregonara et al., 2009, 2011a);

2. **Simultaneous estimation (SIME)**: the AIF is simultaneously estimated with tracer kinetic quantification (Feng et al., 1997; Ogden et al., 2010);

3. **Population-based input function (PBIF)**: individual AIFs are obtained by the mean of several arterial input functions opportunely scaled for individual features (Vriens et al., 2009; Zanotti-Fregonara et al., 2012a);

4. **Venous input function (VIF)**: AIF is derived by the venous blood measured (Syvanen et al., 2006).

The validity of each methods is strongly dependent to the characteristics of tracer kinetics (fast and slow kinetic in particular) as well as to the applied quantification methods. Graphical methods, for example, are not sensitive to the shape of AIF, while depend to its area under curves (AUCs). Since AIF AUC is much more easier to be reproduced respect to its time-course, non invasive input function method have been successfully applied with Logan and Patlak analysis in different tracers (Zanotti-Fregonara et al., 2011a,b, 2012b).

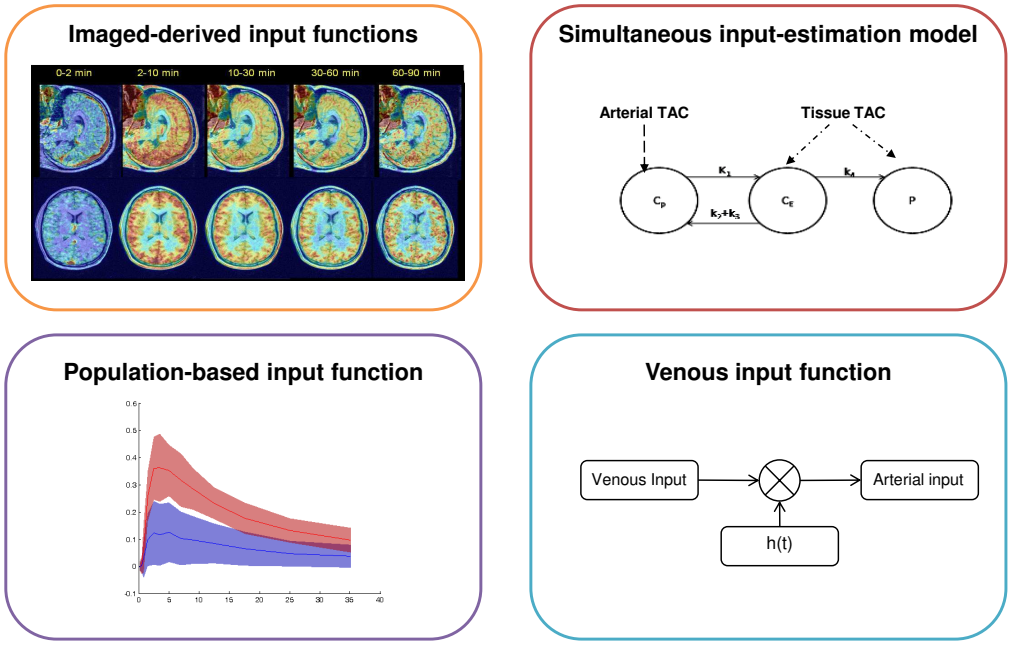

Figure 7.2.1: Alternatives to arterial sampling in PET domain.

On the other hand, compartmental analysis and SA results are strictly dependent to the shape of AIF, and small variations on its time-course can lead to high biased estimates, both micro and macro parameters. Moreover, the application of non invasive methods results in an increased uncertainty of the model parameter estimates, which becomes a problem when groups of subjects have to be statistically compared.
At the moment, it does not exist a valid alternative to arterial sampling able to correctly reproduce its results. Thus, for all the SA-based methods arterial sampling remains a necessity.

7.2.2 Sensitivity of the methods to the algorithm setting

One of the main difficulties in using SA-based methods is related to the choice of the algorithm settings. This step could be particular critical, especially for non modeller users. The problem is even worsened by the lack of a theoretically-based criteria generally valid for all conditions and

all the tracers. The SA-related literature proposes only some *rule of thumbs* that we suggest to be used as a starting point for SA applications, because case by case the correctness of the algorithm setting has to be verified and eventually adapted to the specific characteristics of the applicative context. Hereafter we proposed some guidelines for SA setting definition.

SA BASIS FUNCTION: DISTRIBUTION All the SA methods, standard and filtered versions, require the *a priori* definition of the basis function for the spectral components. The β grid has to be designed to appropriately cover the distribution of all the kinetic components detectable from the PET data, making it more dense where the probability to measure a component is higher. Thus, the choice of β grid can be seen as a way of applying a prior information on the quantification: the more accurate this information is, the better the quantification results are. It has to be noticed that the relationship between basis function and final results is not well-formulated as the Bayesian methods, and since the first application of SA, it has been always preferred to implement a β grid larger than the actual kinetic of the tracer under study (Cunningham and Jones, 1993; Turkheimer et al., 1994).
Also the choice of a logarithmic distributions of the spectral components represents a standard procedure derived principally from the common practice rather than theoretically formulation. In fact, it has been shown in (Cunningham et al., 1998) and in (Turkheimer et al., 2003) that other choices of distribution, like equiangular or orthogonal basis, might improve the performance of SA methods, both standard and filtered versions.
For the particular characteristics of SAIF, we preferred the use of the logarithmic grid. This choice was driven by the necessity to implement a grid whose generation could be simply managed: in fact, since SAIF implements an iterative cycle in which the β grid is involved, the application of more complex methods would create a too computation demanding algorithm. In addition, from our experience of SAIF applications, we did not find any evidence suggesting the necessity to use a different distribu-

tion of βs for SAIF.

SA BASIS FUNCTION: NUMBER OF COMPONENTS The number of grid components appears to be less critical than the choice of their distributions. Different attempts have been done comparing SA results with a different number of βs (Tomasi et al., 2009b; Veronese et al., 2010): the advantages of a more populated grid, especially in term of accuracy and precision of spectral estimates, does not counterbalance the relevant increase of computation time required by the algorithm. On the contrary, a smaller number of components could be critical if not optimally distributed around the true kinetic components. As suggested by Rizzo (2012), 20/30 elements are adequate only when an accurate prior information about the data under analysis is available. If this condition is not verified the final estimates can be strongly biased.
The results obtained in chapter 3 and 4 has confirmed that a grid consisting in 100 βs represents a good trade-off between results precision and algorithm efficiency. We recommend it for both filtered and standard SA applications.

FILTER SETTING Respect to the standard SA, filtered SA methods require also the filtering definition. The filter procedures heavily impact on the final performance of the methods and thus they have to be carefully managed, depending on the peculiarities of SA algorithm under utilization. The difference of using an adequate or not adequate filtering setting is determinant to make the methods reliable or not.
For SAIF, for example, the choice of the filter passband coincides with the assumption of the kinetic interval within which the investigate kinetic can be detected. Outside this interval, all the kinetic components are considered as a mixed of noise, random SA effects and true kinetic components (like blood volume or trapping) and thus removed from the estimated spectrum. For its definition, different strategies based on the *a priori* information available on the particular dataset and tracer under analysis, have been evaluated: for L-[1-^{11}C]Leucine PET data we im-

plemented a simulation approach based on literature kinetic values; for [^{18}F]FDG , instead, we derived the filter passband bounds directly from literature kinetic estimates. Finally, for [^{18}F]FLT , we implemented a hierarchical approach, from ROI to voxel analysis. The last approach used seems to be the most promising, since its definition does not depend on the particular tracer or system under study.
Independently from the strategy used for filter definition, it is always recommended to verify *a posteriori* the assumption correctness. For SAIF, for example, in the particular case in which the spectral components are concentrated at the borders of the filter, the choice of the passband is with high probability too narrow for the kinetic characteristics of the process under study (Fig. 7.2.2). In these cases a new quantification with a wider SAIF passband is highly recommended.

7.3 SAIF: Prons and Cons

All the considerations done for SA can be directly extended to SAIF. Since the method conserves the main proprieties of standard SA, advantages and drawbacks are also maintained. Figure 7.3.1 proposed a schematic summary of SAIF features when applied for PET quantification of irreversible tracers.

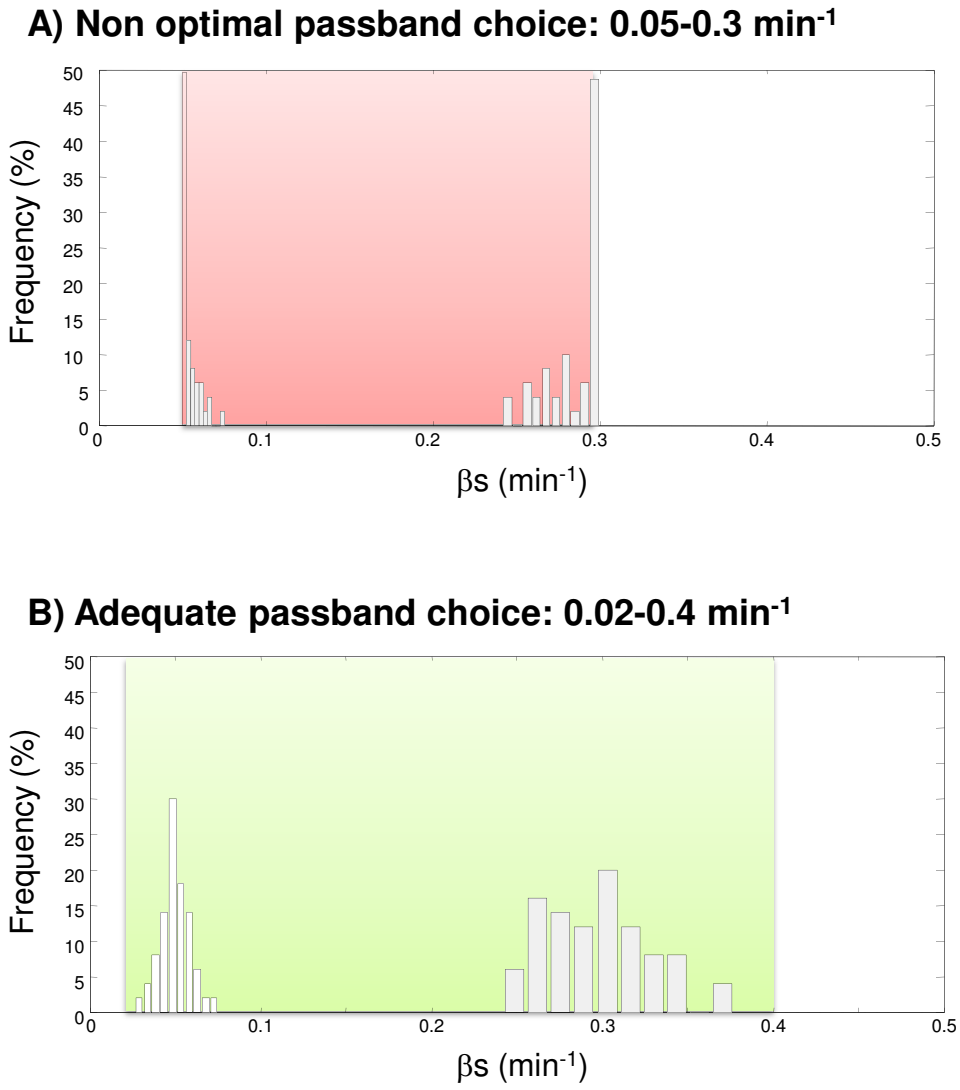

Figure 7.2.2: A posteriori analysis of SAIF kinetic components. Panel A shows the case of over-fitting passband. Most of the components detected coincide with the filter passband bounds. Panel B shows the same situation, with an optimal filter passband. Results refer to SAIF application to the same simulated dataset.

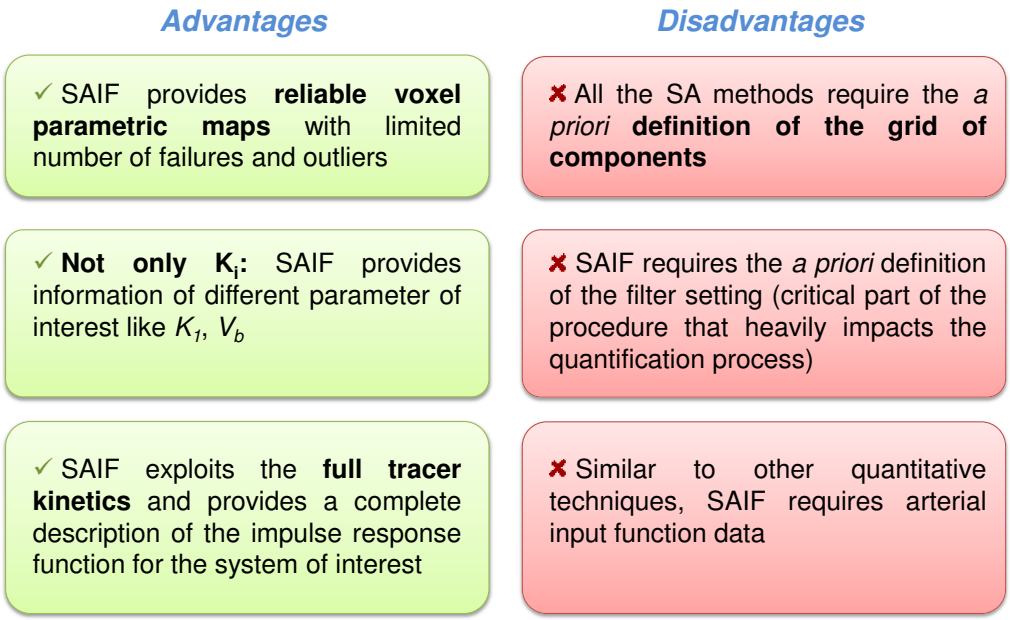

Figure 7.3.1: SAIF prons and cons.

Writing a book is an adventure. To begin with, it is a toy and an amusement; then it becomes a mistress, and then it becomes a master, and then a tyrant. The last phase is that just as you are about to be reconciled to your servitude, you kill the monster, and fling him out to the public.

Winston Churchill

8 Conclusions

Quantification of dynamic PET studies can be performed in several different ways, which mainly differ on the assumptions about the system of interest and on the analysis outcomes. In this work we gave particular attention to Spectral Analysis, a general and flexible quantification method based on minimal model assumptions. Considering the trade-off between outcomes and requirements, Spectral Analysis can be located in an intermediate position between graphical methods and compartmental modelling.

The main limitation of Spectral Analysis is related to its sensitivity to the noise in the data. For this reason we developed SAIF, a filtered Spectral Analysis version for the quantification of irreversible tracers. SAIF was first formalized in the Spectral Analysis context and then validated in different cases of interest, by comparing its results with those provided by the standard Spectral Analysis as well as with the methodological state of the art presented in literature. Results have indicated SAIF as a robust and reliable quantification method, applicable for both region and voxel-wise analysis, in different tracers, anatomical systems

and physiological conditions.
We also developed SAKE, a software tool to make Spectral Analysis friendly usable. SAKE incorporates the main Spectral Analysis methods within a unique environment designed to manage the entire quantification process, from data pre-processing to analysis of results.
Despite the great potential offered by the technique, Spectral Analysis is not so widely applied in PET community, mostly because not completely known and understood. With this work we tried to support its use by offering an exhaustive dissertation of the method advantages and drawbacks, with the ambitious wish to make it more accepted in the PET modelling community.

Bibliography

N. M. Alpert and F. Yuan. A general method of Bayesian estimation for parametric imaging of the brain. *Neuroimage*, 45:1183–1189, May 2009.

P. Bartenstein, S. Asenbaum, A. Catafau, C. Halldin, L. Pilowski, A. Pupi, and K. Tatsch. European Association of Nuclear Medicine procedure guidelines for brain imaging using [(18)F]FDG. *Eur. J. Nucl. Med. Mol. Imaging*, 29(10):P43–48, Oct 2002.

G. Bellani, C. Messa, L. Guerra, E. Spagnolli, G. Foti, N. Patroniti, R. Fumagalli, G. Musch, F. Fazio, and A. Pesenti. Lungs of patients with acute respiratory distress syndrome show diffuse inflammation in normally aerated regions: a [18F]-fluoro-2-deoxy-D-glucose PET/CT study. *Crit. Care Med.*, 37(7):2216–2222, Jul 2009.

G. Bellani, L. Guerra, G. Musch, A. Zanella, N. Patroniti, T. Mauri, C. Messa, and A. Pesenti. Lung regional metabolic activity and gas volume changes induced by tidal ventilation in patients with acute lung injury. *Am. J. Respir. Crit. Care Med.*, 183(9):1193–1199, May 2011.

M. Bentourkia. PET kinetic modeling of 11C-acetate from projections. *Comput Med Imaging Graph*, 27(5):373–379, 2003.

G. R. Bernard. Acute respiratory distress syndrome: a historical perspective. *Am. J. Respir. Crit. Care Med.*, 172(7):798–806, Oct 2005.

A. Bertoldo, P. Vicini, G. Sambuceti, A. A. Lammertsma, O. Parodi, and C. Cobelli. Evaluation of compartmental and spectral analysis models of [(18)F]FDG kinetics for heart and brain studies with PET. *IEEE Trans Biomed Eng*, 45(12):1429–48, 1998.

A. Bertoldo, P. Peltoniemi, V. Oikonen, J. Knuuti, P. Nuutila, and C. Cobelli. Kinetic modeling of [(18)F]FDG in skeletal muscle by PET: a four-compartment five-rate-constant model. *Am. J. Physiol. Endocrinol. Metab.*, 281(3):E524–536, Sep 2001.

A. Bertoldo, J. Price, C. Mathis, S. Mason, D. Holt, C. Kelley, C. Cobelli, and D. E. Kelley. Quantitative assessment of glucose transport in human skeletal muscle: dynamic positron emission tomography imaging of [O-methyl-11C]3-O-methyl-D-glucose. *J. Clin. Endocrinol. Metab.*, 90 (3):1752–1759, Mar 2005.

A. Bertoldo, R. R. Pencek, K. Azuma, J. C. Price, C. Kelley, C. Cobelli, and D. E. Kelley. Interactions between delivery, transport, and phosphorylation of glucose in governing uptake into human skeletal muscle. *Diabetes*, 55(11):3028–3037, Nov 2006.

S. Bishu, K. C. Schmidt, T. Burlin, M. Channing, S. Conant, T. Huang, Z. H. Liu, M. Qin, A. Unterman, Z. Xia, A. Zametkin, P. Herscovitch, and C. B. Smith. Regional rates of cerebral protein synthesis measured with L-[1-^{11}C]leucine and PET in conscious, young adult men: normal values, variability, and reproducibility. *J Cereb Blood Flow Metab*, 28(8): 1502–13, 2008.

S. Bishu, KC Schmidt, and CB Smith. Propofol anesthesia does not alter the regional rates of cerebral protein synthesis measured with L-[1-^{11}C]leucine and PET in healthy male subjects. *J Cereb Blood Flow Metab*, 2009.

R. Boellaard, N. C. Krak, O. S. Hoekstra, and A. A. Lammertsma. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J. Nucl. Med.*, 45(9):1519–1527, Sep 2004.

R. Carson, W. Barker, J. S. Liow, and C. Jonhson. Design of a motion-compensation osem list-mode algorithm for resolution-recovery reconstruction for the hrrt. *IEEE Trans Nucl Sci*, pages 3281–3285, 2003.

R.E. Carson, C. Cobelli, and L. Filkestein. *The Mathematical Modeling of Metabolic and Endocrine Systems*. Wiley, New York, 1983.

K. Chen, D. Bandy, E. Reiman, S. C. Huang, M. Lawson, D. Feng, L. S. Yun, and A. Palant. Noninvasive quantification of the cerebral metabolic rate for glucose using positron emission tomography, 18F-fluoro-2-deoxyglucose, the Patlak method, and an image-derived input function. *J. Cereb. Blood Flow Metab.*, 18(7):716–723, Jul 1998.

R. C. Collins, N. Nandi, C. B. Smith, and L. Sokoloff. Focal seizures inhibit brain protein synthesis. *Trans Am Neurol Assoc*, 105:43–6, 1980.

Cramer and Herald. *Mathematical Methods of Statistics*. Princeton University Press, Princeton, NJ, 1946.

R. A. Cunha, S. Ferre, J. M. Vaugeois, and J. F. Chen. Potential therapeutic interest of adenosine A2A receptors in psychiatric disorders. *Curr. Pharm. Des.*, 14(15):1512–1524, 2008.

V. J. Cunningham and T. Jones. Spectral analysis of dynamic PET studies. *J Cereb Blood Flow Metab*, 13(1):15–23, 1993.

V.J. Cunningham, J. Ashburner, H. Byrne, and T. Jones. Use of spectral analysis to obtain parametric images from dynamic pet studies. In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in Brain PET*. Elsevier, Amsterdam, 1993.

V.J. Cunningham, R.N Gunn, H. Byrne, and J.C. Matthews. Suppression of noise artifacts in spectral analysis of dynamic pet data. In *Quantitative Functional Brain Imaging with Positron Emission Tomography*. Accademic Press, San Diego, 1998.

N. de Prost, M. R. Tucci, and M. F. Melo. Assessment of lung inflammation with 18F-FDG PET during acute lung injury. *AJR Am J Roentgenol*, 195(2):292–300, Aug 2010.

J. J. DiStefano. Optimized blood sampling protocols and sequential design of kinetic experiments. *Am. J. Physiol.*, 240(5):R259–265, May 1981.

A. S. Dittrich, T. Winkler, T. Wellman, N. de Prost, G. Musch, R. S. Harris, and M. F. Vidal Melo. Modeling (18)F-FDG Kinetics during Acute Lung Injury: Experimental Data and Estimation Errors. *PLoS ONE*, 7(10): e47588, 2012.

M. Dowsett, C. Archer, L. Assersohn, R. K. Gregory, P. A. Ellis, J. Salter, J. Chang, P. Mainwaring, I. Boeddinghaus, S. R. Johnston, T. J. Powles, and I. E. Smith. Clinical studies of apoptosis and proliferation in breast cancer. *Endocr. Relat. Cancer*, 6(1):25–28, Mar 1999.

P. Edison, D. J. Brooks, F. E. Turkheimer, H. A. Archer, and R. Hinz. Strategies for the generation of parametric images of [11C]PIB with plasma input functions considering discriminations and reproducibility. *Neuroimage*, 48(2):329–338, Nov 2009.

B Efron. Bootstrap methods: another look at the jackknife. *Ann. Statistic*, 1979.

D. Feng, S. C. Huang, Z. Z. Wang, and D. Ho. An unbiased parametric imaging algorithm for nonuniformly sampled biomedical system parameter estimation. *IEEE Trans Med Imaging*, 15:512–518, 1996.

D. Feng, K. P. Wong, C. M. Wu, and W. C. Siu. A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: theory and simulation study. *IEEE Trans Inf Technol Biomed*, 1(4):243–254, Dec 1997.

K. Fuxe, S. Ferre, S. Genedani, R. Franco, and L. F. Agnati. Adenosine receptor-dopamine receptor interactions in the basal ganglia and their relevance for brain function. *Physiol. Behav.*, 92(1-2):210–217, Sep 2007.

I. B. Galazzo, S. K. Bose, A. F. Ramlackhansingh, I. Ahmed, N. Pavese, G. Rizzo, D. J. Brooks, F. E. Turkheimer, and A. Bertoldo. Kinetic modeling of the adenosine A2A subtype receptor radioligand [11C]SCH442416 in humans. *Neuroimage*, 52:S178, Aug 2010.

Keith Godfrey. *Compartmental Models and Their Application*. Accademic Press, London, 1982.

K. R. Gray, K. B. Contractor, L. M. Kenny, A. Al-Nahhas, S. Shousha, J. Stebbing, H. S. Wasan, R. C. Coombes, E. O. Aboagye, F. E. Turkheimer, and L. Rosso. Kinetic filtering of [(18)F]Fluorothymidine in positron emission tomography studies. *Phys Med Biol*, 55(3):695–709, Feb 2010.

E. Grecchi. *Quantification of lung glucose metabolism with positron emission tomography in patients with Acute Lung Injury*. Padova University, 2011.

G. Grunder, T. Siessmeier, C. Lange-Asschenfeldt, I. Vernaleken, H. G. Buchholz, P. Stoeter, A. Drzezga, H. Luddens, F. Rosch, and P. Bartenstein. [18F]Fluoroethylflumazenil: a novel tracer for PET imaging of human benzodiazepine receptors. *Eur J Nucl Med*, 28(10):1463–1470, Oct 2001.

R. N. Gunn, A. A. Lammertsma, S. P. Hume, and V. J. Cunningham. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *Neuroimage*, 6(4):279–87, 1997.

R. N. Gunn, S. R. Gunn, and V. J. Cunningham. Positron emission tomography compartmental models. *J. Cereb. Blood Flow Metab.*, 21:635–652, Jun 2001.

R. N. Gunn, S. R. Gunn, F. E. Turkheimer, J. A. Aston, and V. J. Cunningham. Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling. *J Cereb Blood Flow Metab*, 22(12): 1425–39, 2002.

L. M. Hamberg, G. J. Hunter, N. M. Alpert, N. C. Choi, J. W. Babich, and A. J. Fischman. The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification? *J. Nucl. Med.*, 35(8):1308–1312, Aug 1994.

A. Hammers, R. Allom, M. J. Koepp, S. L. Free, R. Myers, L. Lemieux, T. N. Mitchell, D. J. Brooks, and J. S. Duncan. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp*, 19:224–247, Aug 2003.

A. Hammers, M. J. Koepp, D. J. Brooks, and J. S. Duncan. Periventricular white matter flumazenil binding and postoperative outcome in hippocampal sclerosis. *Epilepsia*, 46(6):944–948, Jun 2005.

A. Hammers, M. C. Asselin, F. E. Turkheimer, R. Hinz, S. Osman, G. Hotton, D. J. Brooks, J. S. Duncan, and M. J. Koepp. Balancing bias, reliability, noise properties and the need for parametric maps in quantitative ligand PET: [(11)C]diprenorphine test-retest data. *Neuroimage*, 38:82–94, Oct 2007.

G. Hasko and P. Pacher. A2A receptors in inflammation and injury: lessons learned from transgenic animals. *J. Leukoc. Biol.*, 83(3):447–455, Mar 2008.

I.D. Hinz, R. Grachev, D.L. Cutler, J. Hunter, S. Osman, M. Doder, D.J. Brooks, and V.J. Cunningham. In-vivo quantification of A2A receptors in human brain with [11C]SCH442416 and positron emission tomography. *Molecular Imaging and Biology*, 5:1656, May-June 2003.

Akaike Hirotugu. On a decision procedure for system identification. *Engineering approach to computer control*, 1971.

E. J. Hoffman, S. C. Huang, and M. E. Phelps. Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr*, 3(3):299–308, Jun 1979.

Y. T. Hong and T. D. Fryer. Kinetic modelling using basis functions derived from two-tissue compartmental models with a plasma input

function: general principle and application to [18F]fluorodeoxyglucose positron emission tomography. *Neuroimage*, 51:164–172, May 2010.

S. C. Huang. Anatomy of SUV. Standardized uptake value. *Nucl. Med. Biol.*, 27(7):643–646, Oct 2000.

H. Iida, S. Higano, N. Tomura, F. Shishido, I. Kanno, S. Miura, M. Murakami, K. Takahashi, H. Sasaki, and K. Uemura. Evaluation of regional differences of tracer appearance time in cerebral tissues using [^{15}O] water and dynamic positron emission tomography. *J Cereb Blood Flow Metab*, 8(2):285–8, 1988.

M. C. Ingvar, P. Maeder, L. Sokoloff, and C. B. Smith. Effects of ageing on local rates of cerebral protein synthesis in sprague-dawley rats. *Brain*, 108 (Pt 1):155–70, 1985.

R. B. Innis, V. J. Cunningham, J. Delforge, M. Fujita, A. Gjedde, R. N. Gunn, J. Holden, S. Houle, S. C. Huang, M. Ichise, H. Iida, H. Ito, Y. Kimura, R. A. Koeppe, G. M. Knudsen, J. Knuuti, A. A. Lammertsma, M. Laruelle, J. Logan, R. P. Maguire, M. A. Mintun, E. D. Morris, R. Parsey, J. C. Price, M. Slifstein, V. Sossi, T. Suhara, J. R. Votaw, D. F. Wong, and R. E. Carson. Consensus nomenclature for in vivo imaging of reversibly binding radioligands. *J Cereb Blood Flow Metab*, 27(9):1533–1539, 2007. doi: 10.1038/sj.jcbfm.9600493.

C. R. Jiang, J. A. Aston, and J. L. Wang. Smoothing dynamic positron emission tomography time courses using functional principal components. *Neuroimage*, 47(1):184–193, Aug 2009.

L. Kenny, R. C. Coombes, D. M. Vigushin, A. Al-Nahhas, S. Shousha, and E. O. Aboagye. Imaging early changes in proliferation at 1 week post chemotherapy: a pilot study in breast cancer patients with 3'-deoxy-3'-[18F]fluorothymidine positron emission tomography. *Eur. J. Nucl. Med. Mol. Imaging*, 34(9):1339–1347, Sep 2007.

L. M. Kenny, D. M. Vigushin, A. Al-Nahhas, S. Osman, S. K. Luthra, S. Shousha, R. C. Coombes, and E. O. Aboagye. Quantification of cellular proliferation in tumor and normal tissues of patients with breast cancer by [18F]fluorothymidine-positron emission tomography imaging: evaluation of analytical methods. *Cancer Res.*, 65(21):10104–10112, Nov 2005.

C. K. Kim, N. C. Gupta, B. Chandramouli, and A. Alavi. Standardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J. Nucl. Med.*, 35(1):164–167, Jan 1994.

K. Kubota. Lung tumor imaging by positron emission tomography using C-11 L-methionine. *J. Nucl. Med.*, 26(1):37–42, Jan 1985.

K. Lehtio, V. Oikonen, S. Nyman, T. Gronroos, A. Roivainen, O. Eskola, and H. Minn. Quantifying tumour hypoxia with fluorine-18 fluoroerythronitroimidazole ([18F]FETNIM) and PET using the tumour to plasma ratio. *Eur. J. Nucl. Med. Mol. Imaging*, 30(1):101–108, Jan 2003.

J. Logan, J. S. Fowler, N. D. Volkow, A. P. Wolf, S. L. Dewey, D. J. Schlyer, R. R. MacGregor, R. Hitzemann, B. Bendriem, and S. J. Gatley. Graphical analysis of reversible radioligand binding from time-activity measurements applied to [N-11C-methyl]-(-)-cocaine PET studies in human subjects. *J. Cereb. Blood Flow Metab.*, 10:740–747, Sep 1990.

R. C. Marshall, P. Powers-Risius, B. W. Reutter, J. P. O'Neil, M. La Belle, R. H. Huesman, and H. F. VanBrocklin. Kinetic analysis of 18F-fluorodihydrorotenone as a deposited myocardial flow tracer: comparison to 201Tl. *J. Nucl. Med.*, 45(11):1950–1959, Nov 2004.

M. A. Matthay. Treatment of acute lung injury: clinical and experimental studies. *Proc Am Thorac Soc*, 5(3):297–299, Apr 2008.

B. M. Mazoyer, R. H. Huesman, T. F. Budinger, and B. L. Knittel. Dynamic PET data analysis. *J Comput Assist Tomogr*, 10:645–653, 1986.

I. Miederer, S. I. Ziegler, C. Liedtke, M. E. Spilker, M. Miederer, T. Sprenger, K. J. Wagner, A. Drzezga, and H. Boecker. Kinetic modelling of [11C]flumazenil using data-driven methods. *Eur. J. Nucl. Med. Mol. Imaging*, 36(4):659–670, Apr 2009.

J. Mihailovic, S.J. Goldsmith, and R.P. Killeen. *FDG PET/CT in Clinical Oncology*. Springer, Berlin, 2012.

T. Mihara, A. Noda, H. Arai, K. Mihara, A. Iwashita, Y. Murakami, T. Matsuya, S. Miyoshi, S. Nishimura, and N. Matsuoka. Brain adenosine A2A receptor occupancy by a novel A1/A2A receptor antagonist, ASP5854, in rhesus monkeys: relationship to anticataleptic effect. *J. Nucl. Med.*, 49(7):1183–1188, Jul 2008.

S. Mitkovski, V. L. Villemagne, K. E. Novakovic, G. O'Keefe, H. Tochon-Danguy, R. S. Mulligan, K. L. Dickinson, T. Saunder, M. C. Gregoire, M. Bottlaender, F. Dolle, and C. C. Rowe. Simplified quantification of nicotinic receptors with 2[18F]F-A-85380 PET. *Nucl. Med. Biol.*, 32(6): 585–591, Aug 2005.

T. Miyauchi and R. L. Wahl. Regional 2-[18F]fluoro-2-deoxy-D-glucose uptake varies in normal lung. *Eur J Nucl Med*, 23(5):517–523, May 1996.

R. M. Moresco, S. Todde, S. Belloli, P. Simonelli, A. Panzacchi, M. Rigamonti, M. Galli-Kienle, and F. Fazio. In vivo imaging of adenosine A2A receptors in rat and primate brain using [11C]SCH442416. *Eur J Nucl Med Mol Imaging*, 32(4):405–413, 2005. doi: 10.1007/s00259-004-1688-5.

J. F. Myers, L. Rosso, B. J. Watson, S. J. Wilson, N. J. Kalk, N. Clementi, D. J. Brooks, D. J. Nutt, F. E. Turkheimer, and A. R. Lingford-Hughes. Characterisation of the contribution of the GABA-benzodiazepine 1 receptor subtype to [(11)C]Ro15-4513 PET images. *J. Cereb. Blood Flow Metab.*, 32(4):731–744, Apr 2012.

R. T. Ogden. Estimation of kinetic parameters in graphical analysis of PET imaging data. *Stat Med*, 22:3557–3568, Nov 2003.

R. T. Ogden, F. Zanderigo, S. Choy, J. J. Mann, and R. V. Parsey. Simultaneous estimation of input functions: an empirical study. *J. Cereb. Blood Flow Metab.*, 30(4):816–826, Apr 2010.

D. M. Parkin, P. Pisani, and J. Ferlay. Global cancer statistics. *CA Cancer J Clin*, 2008.

C.S. Patlak and R.G. Blasberg. Graphical evaluation o of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab*, 1985.

R. R. Pencek, A. Bertoldo, J. Price, C. Kelley, C. Cobelli, and D. E. Kelley. Dose-responsive insulin regulation of glucose transport in human skeletal muscle. *Am. J. Physiol. Endocrinol. Metab.*, 290(6):E1124–1130, Jun 2006.

J. Y. Peng, J. A. Aston, R. N. Gunn, C. Y. Liou, and J. Ashburner. Dynamic positron emission tomography data-driven analysis using sparse Bayesian learning. *IEEE Trans Med Imaging*, 27(9):1356–1369, Sep 2008.

M. E. Phelps, S. C. Huang, E. J. Hoffman, C. Selin, L. Sokoloff, and D. E. Kuhl. Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: validation of method. *Ann. Neurol.*, 6:371–388, Nov 1979.

M. Qin, J. Kang, T. V. Burlin, C. Jiang, and C. B. Smith. Postadolescent changes in regional cerebral protein synthesis: an in vivo study in the FMR1 null mouse. *J Neurosci*, 25(20):5087–95, 2005.

A. F. Ramlackhansingh, S. K. Bose, I. Ahmed, F. E. Turkheimer, N. Pavese, and D. J. Brooks. Adenosine 2A receptor availability in dyskinetic and nondyskinetic patients with Parkinson disease. *Neurology*, 76:1811–1816, May 2011.

M. P. Richardson, M. J. Koepp, D. J. Brooks, D. R. Fish, and J. S. Duncan. Benzodiazepine receptors in focal epilepsy with cortical dysgenesis: an 11C-flumazenil PET study. *Ann. Neurol.*, 40(2):188–198, Aug 1996.

G. Rizzo. *DEVELOPMENT OF NOVEL COMPUTATIONAL ALGORITHMS FOR QUANTITATIVE VOXEL-WISE FUNCTIONAL BRAIN IMAGING WITH POSITRON EMISSION TOMOGRAPHY*. PhD thesis, Ph.D. School on Information Engineering, Padova University, 2012.

G. Rizzo, F. E. Turkheimer, and A. Bertoldo. Multi-Scale hierarchical approach for parametric mapping: assessment on multi-compartmental models,. *Neuroimage*, 59:2485–2493, Feb 2012a.

G. Rizzo, F. E. Turkheimer, S. Keihaninejad, S. K. Bose, A. Hammers, and A. Bertoldo. Multi-Scale hierarchical generation of PET parametric maps: application and testing on a [11C]DPN study. *Neuroimage*, 59 (3):2485–2493, Feb 2012b.

G. Rizzo, M. Veronese, and A. Bertoldo. Spectral Analysis Kinetic Estimation: an integrated tool for quantification of PET data. *World Molecular Imaging Congress 2012*, March 2012c.

S. N. Schiffmann, G. Fisone, R. Moresco, R. A. Cunha, and S. Ferre. Adenosine A2A receptors and basal ganglia physiology. *Prog. Neurobiol.*, 83:277–292, Dec 2007.

K. Schmidt. Which linear compartmental systems can be analyzed by spectral analysis of PET output data summed over all compartments? *J Cereb Blood Flow Metab*, 19(5):560–9, 1999.

K. C. Schmidt and F. E. Turkheimer. Kinetic modeling in positron emission tomography. *Q J Nucl Med*, 46:70–85, Mar 2002.

K. C. Schmidt, M. P. Cook, M. Qin, J. Kang, T. V. Burlin, and C. B. Smith. Measurement of regional rates of cerebral protein synthesis with L-[1-^{11}C]leucine and PET with correction for recycling of tissue amino acids: I. kinetic modeling approach. *J Cereb Blood Flow Metab*, 25(5):617–28, 2005.

T. Schroeder, M. F. Vidal Melo, G. Musch, R. S. Harris, J. G. Venegas, and T. Winkler. Modeling pulmonary kinetics of 2-deoxy-2-[18F]fluoro-D-glucose during acute lung injury. *Acad Radiol*, 15(6):763–775, Jun 2008.

Health Quality Ontario Service. Positron emission tomography for the assessment of myocardial viability: an evidence-based analysis. *Ont Health Technol Assess Ser*, 5(16):1–167, 2005.

M. Slifstein and M. Laruelle. Models and methods for derivation of in vivo neuroreceptor parameters with PET and SPECT reversible radiotracers. *Nucl. Med. Biol.*, 28:595–608, Jul 2001.

C. B. Smith and J. Kang. Cerebral protein synthesis in a genetic mouse model of phenylketonuria. *Proc Natl Acad Sci U S A*, 97(20):11014–9, 2000.

C. B. Smith, K. C. Schmidt, M. Qin, T. V. Burlin, M. P. Cook, J. Kang, R. C. Saunders, J. D. Bacher, R. E. Carson, M. A. Channing, W. C. Eckelman, P. Herscovitch, P. Laverman, and B. K. Vuong. Measurement of regional rates of cerebral protein synthesis with L-[1-^{11}C]leucine and pet with correction for recycling of tissue amino acids: II. validation in rhesus monkeys. *J Cereb Blood Flow Metab*, 25(5):629–40, 2005.

L. Sokoloff, M. Reivich, C. Kennedy, M. H. Des Rosiers, C. S. Patlak, K. D. Pettigrew, O. Sakurada, and M. Shinohara. The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J. Neurochem.*, 28:897–916, May 1977.

Y. Sun, G. E. Deibler, J. Jehle, J. Macedonia, I. Dumont, T. Dang, and C. B. Smith. Rates of local cerebral protein synthesis in the rat during normal postnatal development. *Am J Physiol*, 268(2 Pt 2):R549–61, 1995.

S. K. Sundaram, O. Muzik, D. C. Chugani, F. Mu, T. J. Mangner, and H. T. Chugani. Quantification of protein synthesis in the human brain using

L-[1-^{11}C]-leucine PET: incorporation of factors for large neutral amino acids in plasma and for amino acids recycled from tissue. *J Nucl Med*, 47(11):1787–95, 2006.

S. Syvanen, G. Blomquist, L. Appel, M. Hammarlund-Udenaes, B. Langstrom, and M. Bergstrom. Predicting brain concentrations of drug using positron emission tomography and venous input: modeling of arterial-venous concentration differences. *Eur. J. Clin. Pharmacol.*, 62(10):839–848, Oct 2006.

M. Takodoro, A.K.P. Jones, Cunningham V.J., D. Sashin, S. Grootoonk, , J. Ashburner, and T. Jones. Parametric images of ^{11}c-diprenorphine binding using spectral analysis of dynamic PET images acquired in 3d. In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in Brain PET*. Elsevier, Amsterdam, 1993.

P. Therasse, S. G. Arbuck, E. A. Eisenhauer, J. Wanders, R. S. Kaplan, L. Rubinstein, J. Verweij, M. Van Glabbeke, A. T. van Oosterom, M. C. Christian, and S. G. Gwyther. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl. Cancer Inst.*, 92(3):205–216, Feb 2000.

J. A. Thie. Understanding the standardized uptake value, its methods, and implications for usage. *J. Nucl. Med.*, 45(9):1431–1434, Sep 2004.

S. Todde, R. M. Moresco, P. Simonelli, P. G. Baraldi, B. Cacciari, G. Spalluto, K. Varani, A. Monopoli, M. Matarrese, A. Carpinelli, F. Magni, M. G. Kienle, and F. Fazio. Design, radiosynthesis, and biodistribution of a new potent and selective ligand for in vivo imaging of the adenosine A(2A) receptor system using positron emission tomography. *J. Med. Chem.*, 43:4359–4362, Nov 2000.

G. Tomasi, A. Bertoldo, and C. Cobelli. PET parametric imaging improved by global-two-stage method. *Ann Biomed Eng*, 37:419–427, Feb 2009a.

G. Tomasi, A. Bertoldo, and K. Schmidt. Voxel-based estimation of kinetic model parameters of the L-[1-^{11}C]leucine PET method for determination of regional rates of cerebral protein synthesis: Validation and comparison with region-of-interest based methods. *Submitted to Journal of Cerebral Blood Flow Metabolism*, 2009b.

G. Tomasi, A. Bertoldo, C. Cobelli, N. Pavese, Y. F. Tai, A. Hammers, and F. E. Turkheimer. Global-two-stage filtering of clinical PET parametric maps: application to [(11)C]-(R)-PK11195. *Neuroimage*, 55:942–953, Apr 2011a.

G. Tomasi, L. Kenny, F. Mauri, F. Turkheimer, and E. O. Aboagye. Quantification of receptor-ligand binding with [F]fluciclatide in metastatic breast cancer patients. *Eur. J. Nucl. Med. Mol. Imaging*, 38(12):2186–2197, Dec 2011b.

G. Tomasi, S. Kimberley, L. Rosso, E. Aboagye, and F. Turkheimer. Double-input compartmental modeling and spectral analysis for the quantification of positron emission tomography data in oncology. *Phys Med Biol*, 57(7):1889–1906, Apr 2012.

F. E. Turkheimer, R. M. Moresco, G. Lucignani, L. Sokoloff, F. Fazio, and K. Schmidt. The use of spectral analysis to determine regional cerebral glucose utilization with positron emission tomography and [18F]fluorodeoxyglucose: theory, implementation, and optimization procedures. *J Cereb Blood Flow Metab*, 14(3):406–22, 1994.

F. E. Turkheimer, L. Sokoloff, A. Bertoldo, G. Lucignani, M. Reivich, J. L. Jaggi, and K. Schmidt. Estimation of component and parameter distributions in spectral analysis. *J Cereb Blood Flow Metab*, 18(11):1211–22, 1998.

F. E. Turkheimer, M. Brett, D. Visvikis, and V. J. Cunningham. Multiresolution analysis of emission tomography images in the wavelet domain. *J Cereb Blood Flow Metab*, 19(11):1189–208, 1999.

F. E. Turkheimer, R. B. Banati, D. Visvikis, J. A. Aston, R. N. Gunn, and V. J. Cunningham. Modeling dynamic PET-SPECT studies in the wavelet domain. *J Cereb Blood Flow Metab*, 20(5):879–93, 2000.

F. E. Turkheimer, R. Hinz, R. N. Gunn, J. A. Aston, S. R. Gunn, and V. J. Cunningham. Rank-shaping regularization of exponential spectral analysis for application to functional parametric mapping. *Phys Med Biol*, 48(23):3819–41, 2003.

F. E. Turkheimer, P. Edison, N. Pavese, F. Roncaroli, A. N. Anderson, A. Hammers, A. Gerhard, R. Hinz, Y. F. Tai, and D. J. Brooks. Reference and target region modeling of [11C]-(R)-PK11195 brain studies. *J. Nucl. Med.*, 48(1):158–167, Jan 2007.

F. E. Turkheimer, S. Selvaraj, R. Hinz, V. Murthy, Z. Bhagwagar, P. Grasby, O. Howes, L. Rosso, and S. K. Bose. Quantification of ligand PET studies using a reference region with a displaceable fraction: application to occupancy studies with [(11)C]-DASB as an example. *J. Cereb. Blood Flow Metab.*, 32(1):70–80, Jan 2012.

M. Veronese, A. Bertoldo, S. Bishu, A. Unterman, G. Tomasi, C. B. Smith, and K. C. Schmidt. A spectral analysis approach for determination of regional rates of cerebral protein synthesis with the L-[1-(11)C]leucine PET method. *J. Cereb. Blood Flow Metab.*, 30:1460–1476, Aug 2010.

M. Veronese, G. Rizzo, and A. Bertoldo. Spectral Analysis Kinetic Estimation: an integrated tool for quantification of PET data. *IXth International Symposium on Functional Neuroreceptor Mapping of the Living Brain*, 2012a.

M. Veronese, K. C. Schmidt, C. B. Smith, and A. Bertoldo. Use of spectral analysis with iterative filter for voxelwise determination of regional rates of cerebral protein synthesis with L-[1-11C]leucine PET. *J. Cereb. Blood Flow Metab.*, 32(6):1073–1085, Jun 2012b.

D. Vriens, L. F. de Geus-Oei, W. J. Oyen, and E. P. Visser. A curve-fitting approach to estimate the arterial plasma input function for the assessment of glucose metabolic rate and response to treatment. *J. Nucl. Med.*, 50(12):1933–1939, Dec 2009.

R. L. Wahl, K. Zasadny, M. Helvie, G. D. Hutchins, B. Weber, and R. Cody. Metabolic monitoring of breast cancer chemohormonotherapy using positron emission tomography: initial evaluation. *J. Clin. Oncol.*, 11 (11):2101–2111, Nov 1993.

R. A. Weeks, V. J. Cunningham, P. Piccini, S. Waters, A. E. Harding, and D. J. Brooks. 11C-diprenorphine binding in Huntington's disease: a comparison of region of interest analysis with statistical parametric mapping. *J. Cereb. Blood Flow Metab.*, 17(9):943–949, Sep 1997.

P. Wells, E. Aboagye, R. N. Gunn, S. Osman, A. V. Boddy, G. A. Taylor, I. Rafi, A. N. Hughes, A. H. Calvert, P. M. Price, and D. R. Newell. 2-[11C]thymidine positron emission tomography as an indicator of thymidylate synthase inhibition in patients treated with AG337. *J. Natl. Cancer Inst.*, 95(9):675–682, May 2003.

H. J. Wester, F. Willoch, T. R. Tolle, F. Munz, M. Herz, I. Oye, J. Schadrack, M. Schwaiger, and P. Bartenstein. 6-O-(2-[18F]fluoroethyl)-6-O-desmethyldiprenorphine ([18F]DPN): synthesis, biologic evaluation, and comparison with [11C]DPN in humans. *J. Nucl. Med.*, 41(7):1279–1286, Jul 2000.

R. Widmann, T. Kuroiwa, P. Bonnekoh, and K. A. Hossmann. [^{14}C]leucine incorporation into brain proteins in gerbils after transient ischemia: relationship to selective vulnerability of hippocampus. *J Neurochem*, 56 (3):789–96, 1991.

R. Widmann, M. Kocher, R. I. Ernestus, and K. A. Hossmann. Biochemical and autoradiographical determination of protein synthesis in experimental brain tumors of rats. *J Neurochem*, 59(1):18–25, 1992.

K Wienhard. The ECAT HRRT: Performance and first clinical application of the new high resolution research tomograph. *IEEE Trans Nucl Sci*, pages 104–110, 2002.

F. Willoch, F. Schindler, H. J. Wester, M. Empl, A. Straube, M. Schwaiger, B. Conrad, and T. R. Tolle. Central poststroke pain and reduced opioid receptor binding within pain processing circuitries: a [11C]diprenorphine PET study. *Pain*, 108(3):213–220, Apr 2004.

Y. Wu and R. E. Carson. Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging. *J Cereb Blood Flow Metab*, 22(12):1440–52, 2002.

P. Zanotti-Fregonara, e. l. M. Fadaili, R. Maroy, C. Comtat, A. Souloumiac, S. Jan, M. J. Ribeiro, V. Gaura, A. Bar-Hen, and R. Trebossen. Comparison of eight methods for the estimation of the image-derived input function in dynamic [(18)F]-FDG PET human brain studies. *J. Cereb. Blood Flow Metab.*, 29(11):1825–1835, Nov 2009.

P. Zanotti-Fregonara, K. Chen, J. S. Liow, M. Fujita, and R. B. Innis. Image-derived input function for brain PET studies: many challenges and few opportunities. *J. Cereb. Blood Flow Metab.*, 31(10):1986–1998, Oct 2011a.

P. Zanotti-Fregonara, J. S. Liow, M. Fujita, E. Dusch, S. S. Zoghbi, E. Luong, R. Boellaard, V. W. Pike, C. Comtat, and R. B. Innis. Image-derived input function for human brain using high resolution PET imaging with [C](R)-rolipram and [C]PBR28. *PLoS ONE*, 6(2):e17056, 2011b.

P. Zanotti-Fregonara, C. S. Hines, S. S. Zoghbi, J. S. Liow, Y. Zhang, V. W. Pike, W. C. Drevets, A. G. Mallinger, C. A. Zarate, M. Fujita, and R. B. Innis. Population-based input function and image-derived input function for [C](R)-rolipram PET imaging: methodology, validation and application to the study of major depressive disorder. *Neuroimage*, 63(3): 1532–1541, Nov 2012a.

P. Zanotti-Fregonara, J. S. Liow, C. Comtat, S. S. Zoghbi, Y. Zhang, V. W. Pike, M. Fujita, and R. B. Innis. Image-derived input function in PET brain studies: blood-based methods are resistant to motion artifacts. *Nucl Med Commun*, 33(9):982–989, Sep 2012b.

A

Voxel-wise quantification of [^{11}C](R)-Rolipram PET studies

A.1 Introduction

[^{11}C](R)-Rolipram is a PET radioligand for the quantification in vivo of phosphodiesterase type IV (PDE4), an enzyme that metabolizes 3′, 5′-cyclic adenosine monophosphate (cAMP). The cAMP cascade plays an important role in major depressive disorder (MDD) and is a potential target for drug development. A recently published study using PET and [^{11}C](R)-Rolipram showed for the first time that PDE4 is downregulated *in vivo* in unmedicated individuals with MDD[1].
A previous study in humans showed that the pharmacokinetics of [^{11}C](R)-Rolipram is well described by a two-tissue four-rate constant compartmental model (2TCM), and that Logan graphical analysis yielded similar results at the region of interest (ROI) level[2]. While the Logan plot allows the easy computation of parametric images for the total distribution volume (V_T), it does not account for the vascular component within the voxel, and generally underestimates the parameter of interest in noisy data[3], especially when applied voxel-wise. Nevertheless, the resulting bias is almost constant, and thus these limitations can be neglected when using Logan plot in statistical comparison between different groups of subjects.

It is important to note, however, that voxel-wise analysis of the whole kinetic of the tracer in tissues may yield the ability to generate quantitatively accurate parametric maps and thus harness the full potential of [^{11}C](R)-Rolipram for the study of MDD and related diseases. This study aimed to assess the best methodology for voxel-wise quantification of [^{11}C](R)-Rolipram in brain PET studies, in order to generate precise and accurate parametric maps of the parameters of interest.

A.2 Materials and Methods

A.2.1 Dataset

The analysis of this study were based on a dataset of 10 healthy subjects. All the subjects underwent a 90-min PET scan following injection of 421$\pm$144 MBq of [^{11}C](R)-Rolipram , complete with metabolite-corrected plasma input functions. T1-weighted MRI was performed as well providing the anatomical information for the region placement. A detailed description of experimental protocol and data pre-processing is reported in [2].

A.2.2 Quantification methods

Several approaches were applied for voxel-wise quantification of [^{11}C](R)-Rolipram PET images. The 2TCM microparameters (K_1, k_2, k_3, k_4, and V_b) were quantified for each voxel TAC by using the Weighted Non-Linear Least Squares (WNLLS). This approach was used based on the assumption that the model that describes tracer kinetics in the ROI can also be extended at the voxel level. The initial parameters were derived from the ROI estimates described in the literature[2]. Due to the weakness of the nonlinear estimator applied at the voxel level (convergence issues, high computational time, sensitivity to initial estimates), we also identified the 2TCM using H-BFM, as implemented by Rizzo and colleagues[4]. H-BFM identifies all the system micro- and macroparameters (i.e. K_1, k_2, k_3, k_4, V_b and V_T) following a hierarchical scheme from ROI to voxel analysis.
As alternative to the model-based solutions, standard Spectral Analysis (ESA) and Rank-shaping Spectral analysis (RS)[5,6] were also applied to measure the volume of distribution (V_T) at the voxel level. First, ROI analysis was performed to set and validate the application of these two methods by comparing their results with those provided by 2TCM applied ROI level. Then the same setting was extended for SA and RS voxel-wise

Table A.2.1: Quantification methods

Method	Type of method	Type of Estimator	Parameter of Interest	Complete kinetic description
WNLLS	Model-based	Non Linear	K_1, k_2, k_3, k_4, V_b and V_T	Yes
H-BFM	Model-based	Linear	K_1, k_2, k_3, k_4, V_b and V_T	Yes
ESA	Data-driven	Linear with nonnegative constraints	K_1, V_b and V_T	Yes
RS	Data-driven	Linear	V_T	Yes
Logan	Graphical Analysis	Linear	V_T	Yes

application. In addition to the methods described above, we also quantified [^{11}C](R)-Rolipram binding with Logan analysis[7]. A summary of the methods applied is reported in Table A.2.1.

A.2.3 Performance indexes

Results obtained at voxel level using 2TCM model solved with Weighted Non-Linear Least Squares (WNLLS) were used as the gold standard.
From the parametric maps obtained with the different methods we discarded those with biologically implausible results; specifically those with a V_T values < 0 and also those with V_T values $>$1.5 (ml/cm^3). For each subject, we therefore only considered the subset of voxels in which all the methods reached convergence, and used the statistical analysis in these groups of voxels.
For each method, we averaged the V_T estimates obtained in the voxels composing each ROI, and considered only the subset of voxels in which all the methods converged. These average values were compared to the average of the V_T estimates obtained with WNLLS in the same voxels in terms of mean relative differences.

A.3 Results

At the ROI level all methods provided comparable V_T results with 2TCM, with high correlation ($R^2 > 0.95$). Logan analysis tended to underestimate V_T, especially in high binding ROIs, while RS showed a constant bias (+9%) probably due to a bias introduced by the method filtering procedure. SA provided the best performance with a computational time x3 lower than 2TCM.
All the parametric maps are presented before correction for outliers in Figure A.3.1. Here, both WNLLS and SA V_T estimates were visually very noisy, while H-BFM, RS, and Logan had only a limited or negligible amount of outlying voxels. WNLLS yielded the poorest image quality, and more than 50% of the voxels had to be eliminated. In particular, 16%

on average of the voxels had to be discarded for non-physiological estimates, 29% for values with a CV $>$ 50%, and 12% for values with an imaginary CV. Moreover, WNLLS required high computational time (up to five times greater than the other methods). Although the V_T estimates in the remaining voxels are considered as the gold standard, these results clearly highlight the necessity of an alternative method for voxel wise quantification.
2TCM-WNLLS applied voxel-wise required high computational time (up to x5 compared to the other methods), and image quality was generally poor. On average, 50% of the voxels were eliminated because their values were physiologically implausible or inaccurate. Although 6% of the voxels had a negative V_T, 2TCM-HBFM showed one of the highest correlations ($R^2 > 0.97$), the best agreement with 2TCM-WNLLS estimates (Figure A.3.2) and the lowest percentage relative difference. Logan plot results presented good correlation but the higher underestimation among the methods. SA performed better than RS (SA: $R^2 > 0.98$; RS: $R^2 > 0.96$) but the image quality was poorer, and 5% of the voxels had to be eliminated. Similarly to ROI results, RS showed a constant bias (-5%) of V_T, probably due to the filtering procedure.

A.4 CONSIDERATIONS

2TCM-WNLLS high computation time and poor image quality point to the necessity of an alternative method for quantification at the voxel level. Compared to the other methods, 2TCM-HBFM showed more accurate V_T quantification with good quality parametric maps and excellent agreement with 2TCM-WNLLS results. Consequently 2TCM-HBFM was selected as the best alternative to the gold standard approach for voxel-wise quantification of [^{11}C](R)-Rolipram data. Among the data-driven methods, SA represented the best approach both at the ROI and voxel level.
In conclusion, different techniques exist to generate parametric images, and each has its own advantages and drawbacks. H-BFM allows the most complete and reliable quantification of the [^{11}C](R)-Rolipram kinetic at the voxel level. If the whole description of the tracer kinetics is not necessary, or if only the V_T estimate is required to compare different populations or in longitudinal studies, then Logan plot represents a valid alternative. Its marginally higher bias is outweighed by the ease of implementation and robustness of the method.

A.5 References

[1] Fujita M, Hines CS, Zoghbi SS, Mallinger AG, Dickstein LP, Liow JS, et al. (2012): Downregulation of brain phosphodiesterase type IV measured with [^{11}C](R)-Rolipram PET in major depressive disorder. Biol Psychiatry. In press.
[2] Zanotti-Fregonara P, Zoghbi SS, Liow JS, Luong E, Boellaard R, Gladding RL, et al. (2011): Kinetic analysis in human brain of [^{11}C](R)-Rolipram , a positron emission tomographic radioligand to image phosphodiesterase 4: a retest study and use of an image-derived input function. Neuroimage. 54:1903-1909.
[3] Slifstein M, Laruelle M (2000): Effects of statistical noise on graphic analysis of PET neuroreceptor studies. J Nucl Med. 41:2083-2088.
[4] Rizzo G, Turkheimer FE, Bose SK, Bertoldo A (2011): Multi-scale hierarchical approach for parametric mapping: assessment on multi-compartmental models. Book of Abstracts, Xth International Conference on Quantification of Brain Function with PET, Barcelona, Spain, 25-28 May 2011.
[5] Cunningham VJ, Jones T (1993): Spectral-Analysis of Dynamic Pet Studies. Journal of Cerebral Blood Flow and Metabolism. 13:15-23.
[6] Turkheimer FE, Hinz R, Gunn RN, Aston JAD, Gunn SR, Cunningham VJ (2003): Rank-shaping regularization of exponential spectral analysis for application to functional parametric mapping. Phys Med Biol. 48:3819-3841.
[7] Logan J, Fowler JS, Volkow ND, Wolf AP, Dewey SL, Schlyer DJ, et al. (1990): Graphical analysis of reversible radioligand binding from time-activity measurements applied to [N-11C-methyl]-(-)cocaine PET studies in human subjects. J Cereb Blood Flow Metab. 10:740-747.

A.6 Project Partnership

This research was conducted in collaboration with P. Zanotti-Fregonara (Molecular Imaging Branch, National Institute of Mental Health). Portions of this work were presented at the IXth International Symposium on Functional Neuroreceptor Mapping of the Living Brain, 9th11th August 2012, Baltimore, USA.

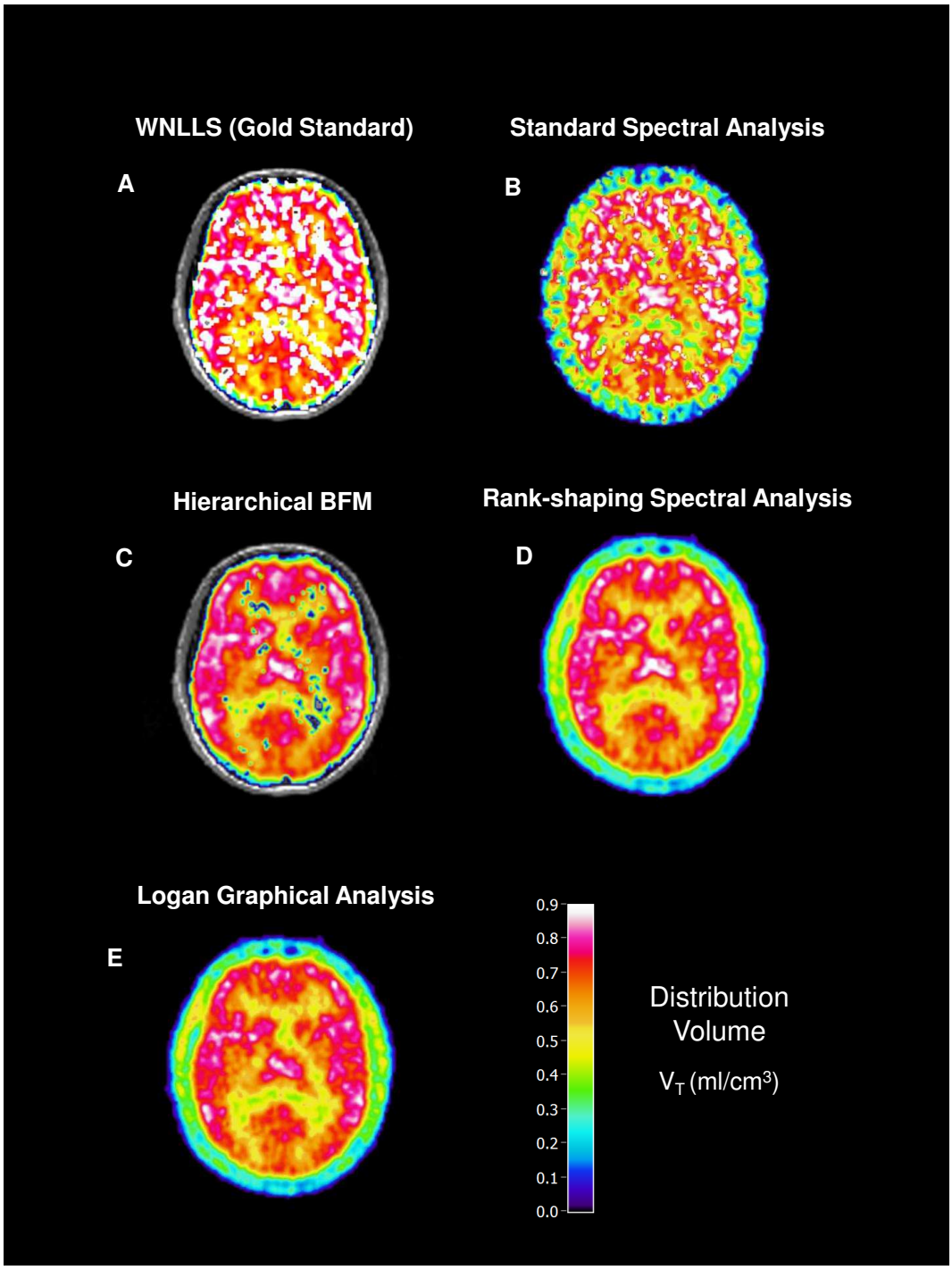

Figure A.3.1: Distribution Volume (V_T) parametric maps obtained with (A) Weighted NonLinear Least Squares (WNLLS) applied to a two-tissue compartmental model, (B) standard Spectral Analysis (ESA), (C) Rank-shaping Spectral Analysis (RS), (D) Hierarchical Basis Function Method (H-BFM) and (E) Logan Graphical Analysis. Results refer to a transaxial slice at the cortical level for a representative subject.

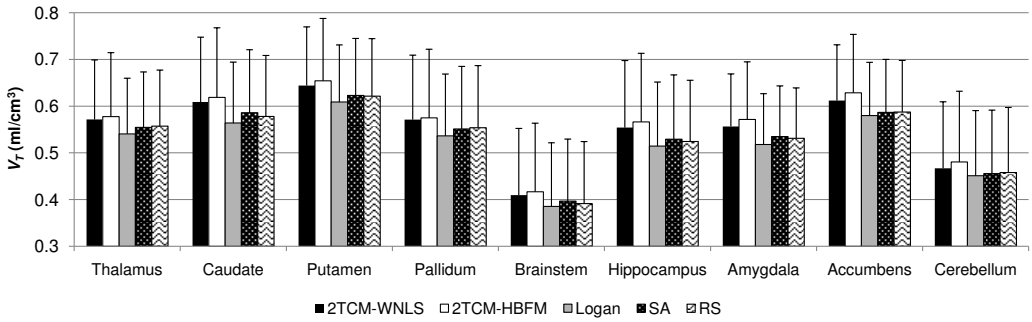

Figure A.3.2: Comparison between 2TCM-WNLLS and other methodologies at the voxel level (variability expressed as between subjects SD).

B

A population-based approach for plasma metabolite correction in PET studies

B.1 Introduction

A notable problem in the quantitative analysis of dynamic PET scans is the accurate estimate of the parent plasma arterial input function due to the presence in plasma of radiolabeled metabolites derived from the metabolism of parent tracer in the liver, kidneys or other parts of the body. When radiolabeled metabolites are found in the plasma in significant amounts their presence has to be accounted for, because it is the concentration of authentic parent tracer that is required for quantification of the tissue data. Estimates of the levels of radiolabeled metabolites in plasma can be assayed directly from samples of blood taken throughout the PET scans by using high-performance liquid chromatography (HPLC). However, for practical and technical reasons the number of metabolite samples is limited (typically to 4-7 for a 1-2 hour scan with ^{11}C or ^{18}F) and they can be noisy particularly later in the scan as counting statistics deteriorate.
Thus, in order to describe the full parent plasma fraction time course, a modelling approach based on the assayed metabolite fractions is applied.

The parent arterial plasma input function is generated by multiplying the measured total plasma radioactivity concentration by the continuous estimate of the parent fraction (PPf). PPf models are usually based on monotonic functional forms including exponentials[1] and sigmoidal[2,3] behaviour. There are several advantages to modelling the fraction of labelled metabolites in plasma: 1) it reduces the effect of measurement noise, 2) it allows the reconstruction of a time-continuous metabolite fraction curve by interpolating few metabolites measures and 3) it permits the extrapolation of metabolite measures when the metabolite fraction can be no longer measured accurately because of low plasma radioactivity levels.
The particular parent fraction model used to correct the plasma time-activity curves for labelled metabolites in PET measurements is chosen to provide the best fit to the data. The choice should be made by integrating this assessment across all subjects and scans within the study so that one model is selected for consistent use with all scans.
The common approach to model-fitting of metabolite samples is based on an individual approach (IND): a parent fraction model for each subject is fitted solely to the metabolite data obtained from that subject. IND for metabolite correction can be applied successfully in a data-rich situation, i.e. in presence of a satisfactorily abundant sampling schedule for each subject. However, in the case of sparse or noisy data, the estimation of the parent fraction may be poor, consequently yielding large errors in the regional tissue parameters of interest following tracer kinetic analysis.
The aforementioned limitations of the IND can be overcome by using Non Linear Mixed Effects (NLMEM[4]). NLMEM is a population-based method for quantitative data analysis and given a model for a system of interest, NLMEM assumes that the set of model parameters, characterized by some physiologically meaningful attributes, can be partitioned into those that do not vary across the population (called fixed effects) and those that do (called random effects). Thus NLMEM easily accommodates both intra-subject as inter-subject variability. Moreover, with NLMEM it is possible to calculate both the population and individual parameter estimates, optimizing all the information derived from all the subject measurements.
In this study, a NLMEM is proposed for modelling the plasma metabolite fraction and generation of the arterial parent plasma input function in dynamic PET scans. Since NLMEM has been developed to provide robust parameter estimates in the case of sparse sampling schedules and/or noisy data, it is merits evaluation in comparison to traditional individ-

ual methods, such as weighted least squares, for the modelling of the parent data.

B.2 Materials and Methods

In order to validate the proposed methodology and to evaluate its impact on arterial input function estimation and on the subsequent tracer kinetic analysis of tissue data, three different PET datasets were considered: [11C]-(+)-PHNO (54 scans), [11C]-PIB (22 scans) and [11C]-DASB (30 scans). For each tracer both simulated and measured data were analyzed by assuming the literature PPf model [5-7] as reference for the analysis. In simulations, the Residual Sum of Squares (RSS) between simulated and model estimated PPf and percentage bias were used as performance indices. In measured data analysis, the mean relative differences between NLMEM and individual results were performed. The impact of metabolite correction was evaluated both at the level of description of parent plasma data and at level of influence on tissue kinetic quantifications.
When the IND approach was used, PPf model parameters were estimated by non-linear weighted residual sum of squares. For NLMEM, the estimates were computed with NONMEM® by assuming log-normal distributions for the PPf parameters and a diagonal covariance matrix to model the inter-subject variability. The same model error description (additive, Gaussian with zero mean and SD defined by the metabolite measures uncertainty) was used in both IND and NLMEM methods.

B.3 Results

Simulations showed that NLMEM provides a more accurate metabolite description than individual approach ($RSS_{pop}/RSSind = 0.24$), particularly in presence of outliers. This NLMEM was shown to be superior even when the simulated metabolite data are equal to the number of PPf model parameters, i.e. data poor scenario, (Figure B.3.1) with a percent difference in the distribution volume (V_T) up to 10%. In measured data analysis, the differences in PPf modelling between NLMEM and IND do not significantly affect neither the AIF reconstruction (Area Under Curve mean relative differences between 0.0% to 3.0%) nor distribution volume (V_T) estimates (absolute mean relative differences between 1% to 4%). This behaviour was found to be consistent in all the analyzed tracers

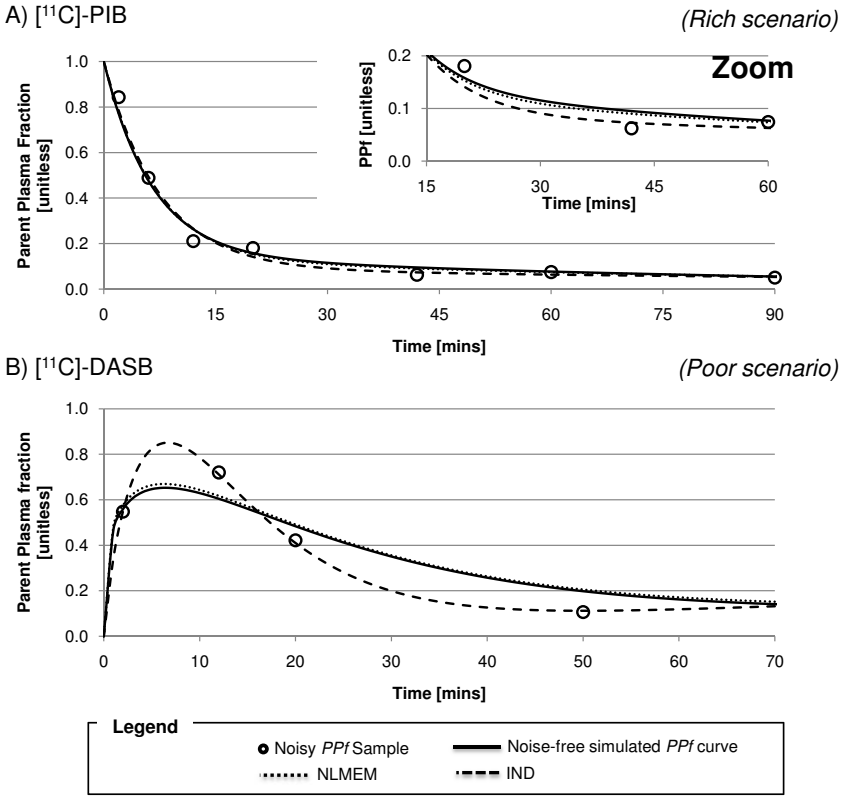

Figure B.3.1: Comparison of NLMEM and IND methods for plasma metabolite correction in simulations. The figure compares the data model fit of IND and NLMEM methods. Panel A refers to [11C]-PIB PPf data in rich scenario conditions. Panel B refers respectively to [11C]-DASB PPf data in poor scenario conditions. For each tracer a representative simulated subject is shown. Circles represent noisy PPf simulated samples; solid lines represent simulated noise-free PPf curves; dashed lines represent the results for IND; dotted lines represent the results for the NLMEM method.

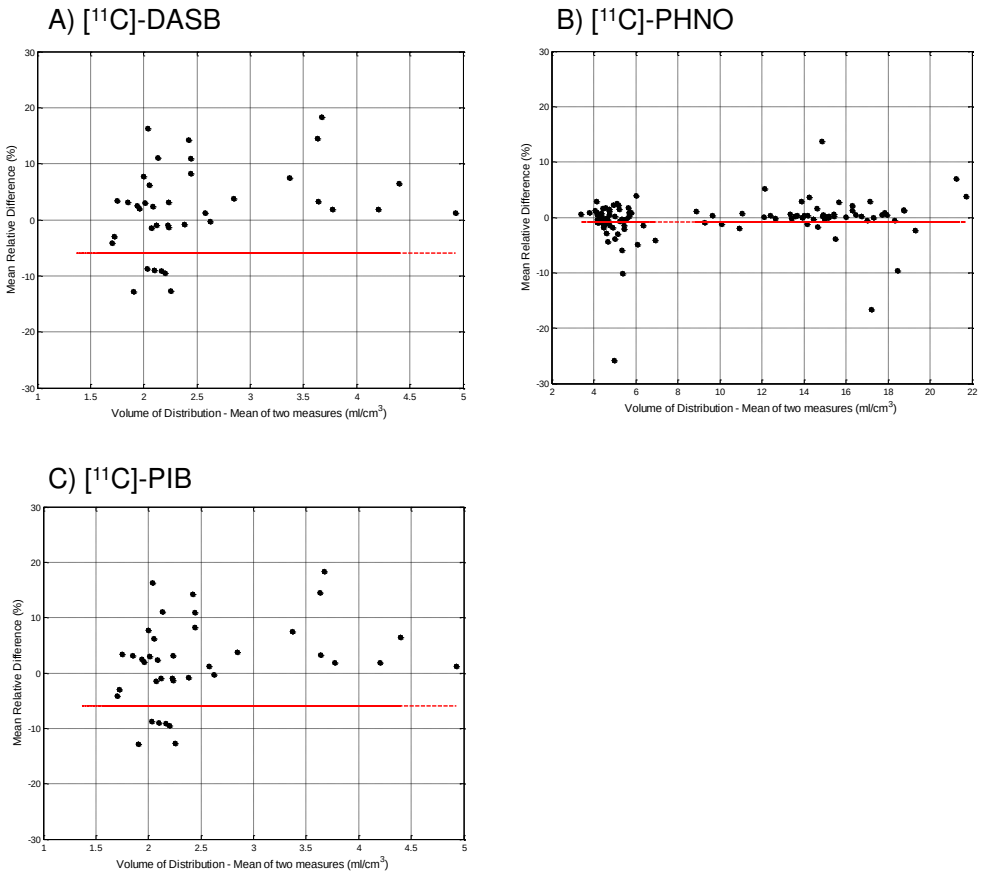

Figure B.3.2: NLMEM vs. IND approaches in Bland-Altman plots. Panels A-C report respectively [11C]-DASB, [11C]-PHNO and [11C]-PIB volume of distribution estimates. Solid lines represent the mean of NLMEM and IND relative differences.

(Figure B.3.2). However, even with real data, there are a few subjects (5%) inside the evaluated populations for which high discrepancy between NLMEM and IND is found.

B.4 CONSIDERATIONS

The application of a NLMEM approach for plasma metabolite correction is shown to offer improvements in performance over the individual approach, particularly in the presence of outliers or when the number of metabolites samples is limited.
For the remaining scenarios, NLMEM performs similarly to the standard individual approach. However, NLMEM requires a population of at least 10 subjects to be viable.

B.5 REFERENCES

[1] Abi-Dargham A, Simpson N, Kegeles L, Parsey R, Hwang DR, Anjilvel S, Zea-Ponce Y, Lombardo I, van Heertum R, Mann JJ, Foged C, Halldin C, Laruelle M. PET Studies of Binding Competition Between Endogenous Dopamine and the D1 Radiotracer [11C]NNC 756. Synapse 1999; 32:93-109.
[2] Gunn RN, Sargent PA, Bench CJ, Rabiner EA, Osman S, Pike VW, Hume SP, Grasby PM, Lammertsma AA. Tracer kinetic modeling of the 5-HT1A receptor ligand [carbonyl-11C]WAY-100635 for PET. Neuroimage 1998; 8: 426-440.
[3] Watabe H, Channing MA, Der MG, Adams HR, Jagoda E, Herscovitch P, Eckelman WC, Carson RE. Kinetic analysis of the 5-HT2A ligand [11C]MDL 100,907. J Cereb Blood Flow Metab 2000; 20:899-909.
[4] Davidian M and Giltinan DM, Nonlinear Models For Repeated Measurement Data, Chapman & Hall/CRC, 1995.
[5] Searle G, Beaver JD, Comley RA, Bani M, Tziortzi A, Slifstein M, Mugnaini M, Griffante C, Wilson AA, Merlo-Pich E, Houle S, Gunn R, Rabiner EA, Laruelle M. Imaging dopamine D3 receptors in the human brain with positron emission tomography, [11C]PHNO, and a selective D3 receptor antagonist. Biol Psychiatry. 2010; 68(4):392-9.
[6] Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, Bergstrm M, Savitcheva I, Huang GF, Estrada S, Ausn B, Debnath ML, Barletta J, Price JC, Sandell J, Lopresti BJ, Wall A, Koivisto P, Antoni G, Mathis CA, Lngstrm B. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Ann Neurol. 2004; 55(3):306-19.

[7] Parsey RV, Ojha A, Ogden RT, Erlandsson K, Kumar D, Landgrebe M, Van Heertum R, Mann JJ. Metabolite considerations in the in vivo quantification of serotonin transporters using 11C-DASB and PET in humans. J Nucl Med. 2006; 47(11):1796-802.

B.6 Project Partnership

This research was conducted in collaboration with prof. Roger Gunn (Imanova Limited, London, UK) and Stefano Zamuner (GSK, London, UK). Portions of this work were presented at XXVth International Symposium on Cerebral Blood Flow, Metabolism and function (Barcellona, 2011) and at Third National Bioingenerring congress (Rome, 2012).

C

Optimal experimental designs in PET receptor occupancy studies

C.1 Introduction

The use of PET imaging has greatly contributed to increase the efficiency of the drug development process of new molecular entities[1]. For neuroscience in particular, PET technique has been widely applied to investigate the ligand-receptor binding in living brain as well as for the determination of plasma concentration/receptor occupancy time-course. Since the knowledge of receptor occupancy and the plasma drug time course are crucial for a rational drug dose selection, PET occupancy studies has proved to be a valuable tool for clinical drug development[2,3].
Despite their high informative content, PET receptor occupancy (PET-RO) studies present some relevant constraints, both ethical and practical. In particular the overall radiation exposure and the high cost of PET experiments, limit the total number of subjects available in the study as well as the and number of scans to perform per subject, preventing the possibility to have dataset with data abundance. Since each scan represents a precious experimental data for the characterization of a drug Pharmacokinetic/Pharmacodynamic (PK/PD) behaviour, rational strategies should be adopted to define the best experiential design.
Optimal experimental designs have been successfully used to increase

the efficiency and minimize cost of the trials by optimizing dose allocation and sampling schedules[4,5]. These techniques allow to increase precision and accuracy of population parameter estimates[6-8] by providing optimized experimental designs which account for the proprieties of the system of interest as well as the subject response variability. All these methods rely on maximization of the information content which can be extract from a particular experiment. They mainly differ on the type of information to maximize and on the applicative requirements. A simple classification can be done distinguishing fixed from adaptive algorithms. Fixed optimal design algorithms use the model description of the investigated system to set the experimental variables by minimizing the estimation error of model parameter estimates. These approaches are simple to be applied, robust and with a very solid theoretical background behind. In theory they represents the best way for the experimental design definition. However some limitations are present. Firstly the application of optimal design algorithms requires detailed information about the model of the investigated system and reliable prior values for its parameters[9-11]. This information is very difficult to be obtained in advance, especially when new chemical entities are studied for the first time in clinical trials. In fact, although the kinetic and dynamic proprieties of a drug can be derived from in vitro or preclinical experiments, a conspicuous level of uncertainty remains when translating this information from animals to humans. Secondly fixed optimal design algorithms provide only theoretic model-based solutions which may not take into account all the practical constraints that an experiment has when run in practice[12-14]. Thus, because the optimization could lead to unfeasible practical designs, it follows that in practice the necessity of comparing different designs performance is as important as to optimize the experiment design. In addition, most of these optimal design techniques planned in advance the entire experimental project without accounting for the results derived by the experiments. This strategy may not be ideal since it is in contrast with the natural way to conduct the PK/PD trials. In dose finding studies, for example, the experiments are commonly managed according to a sequential adaptive procedure: the population subjects are organized in cohorts; the first cohort is then treated with an initial dose (selected on the basis of pre-clinical pharmacokinetic information and early human PK data) and with a measurement sampling grid chosen to appropriately cover the time-window of the drug kinetics (with particular attention to the extremes); afterwards the decision on sample size, dose and measurement times for the second cohort of subjects is derived from the results of the

analysis of the initial data. This procedure is repeated for all the groups of subjects planned in the study. The number of treated subjects, the selection of informative doses and the scan times remain critical issues for a precise and accurate characterization of the PK/PD relationship.
Adaptive optimal designs (AOD) have been proposed for the sequentially optimization of the design in population studies[15,16]. These are educated approaches which combine the information derived from the already performed experiments with that extracted from the model-prediction of the investigated system. This fact requires to organize the population of subjects in cohorts, each one with a different experimental design. Respect to fixed approaches AOD presents some considerable advantages: firstly AOD algorithms reflect more the sequential way with which the experiments are managed in practice, allowing to refine the choice of experimental variables as the experiments are performed and their results analyzed; secondly AOD presents a superior robustness to the prior misspecification which can be corrected step by step with the running of experiments; thirdly, with an infinite number of cohorts, also AOD converge to the optimal design solution.
By considering all these features AOD appears to be more adequate for experimental design definition. In real practice the convenience of using adaptive respect to fixed procedures depends on the specific characteristics of the experiment to which the methods have to be applied. For example, the change of experimental design represents a cost and it cannot be always performed. It can also happen, like in PET occupancy studies, that the number of cohorts in which a study is be organized is limited, counting the advantage of AOD respect to fixed designs. Thus, the type of experimental design strategy to be used should be evaluated case by case, considering all the experimental constraints of the system.
In a recently study it has been demonstrated that adaptive optimal design (AOD) algorithms improve the assessment of drug kinetic time-courses[16], when applied to PET-RO studies. However the value of applying adaptive or non-adaptive optimal design methodologies to PET-RO studies depends on several factors including drug affinity to the target as well as feasibility constrains such as sample size, number of scan per subjects and logistical constrains.
In this work we present a simulation study to explore the potentialities of optimal design algorithms when applied to PET-RO, by evaluating the sensitivity of the results to experimental variables like the scanning times and the patient grouping, as well as misspecified drug kinetic assumptions.

C.2 METHODS

C.2.1 RECEPTOR-TIME COURSE MODEL USING BINDING POTENTIAL

A general representation of PK-occupancy time course model can be described as reported in Figure C.2.1. This modelling is however inapplicable in a real PET drug occupancy study, because only few PET scans per subject (generally <3 scans, one at the baseline and two after the drug administration) can be acquired. Thus a simplified version is required. In our study we considered a binding potential $k_{on} - k_{off}$ model defined as:

$$\frac{dBP(t)}{dt} = k_{off} \cdot BP_0 - [C_p(t) \cdot k_{on} + k_{off}] \cdot BP(t) \tag{C.1}$$

where $BP(t)$ indicates the time dependent binding potential after dosing, BP_0 is the baseline binding potential and $C_p(t)$ the plasma concentration of the drug. k_{on} and k_{on} represent respectively the association and dissociation rates between free brain concentration and receptor. This is equivalent to an assumption that the exchange across the blood brain barrier and the free to non-specifically bound partitioning of the drug in plasma and tissue is sufficiently rapid for the association and dissociation of the specifically bound drug to be considered rate limiting.
From the knowledge of k_{on} and k_{off} it is possible to derive the drug affinity (K_d) to the receptors:

$$K_d = \frac{k_{on}}{k_{off}} \tag{C.2}$$

Generally in a PET-RO study the experimental design is structured as follows: with the baseline scan (scan before drug assumption) BP_0 is derived; with other two PET scans after drug administration k_{on} and k_{off} can be estimated. Since the model represented in Eq. C.1 consists in 3 parameters (BP_0, k_{on} and k_{off}) and only 3 PET scans are available, we are in the case in which the number of experimental measures coincides with the number of model parameters to be estimated. These are severe highly-restrictive data-poor conditions which underline how critical the temporal location of PET scans is.

C.2.2 PRACTICAL ISSUES FOR PET-RO STUDIES

The experimental management of the PET scans in a drug occupancy study follows some empirical rules of thumb defined according to the practical issues in which the study is performed. First of all the use of

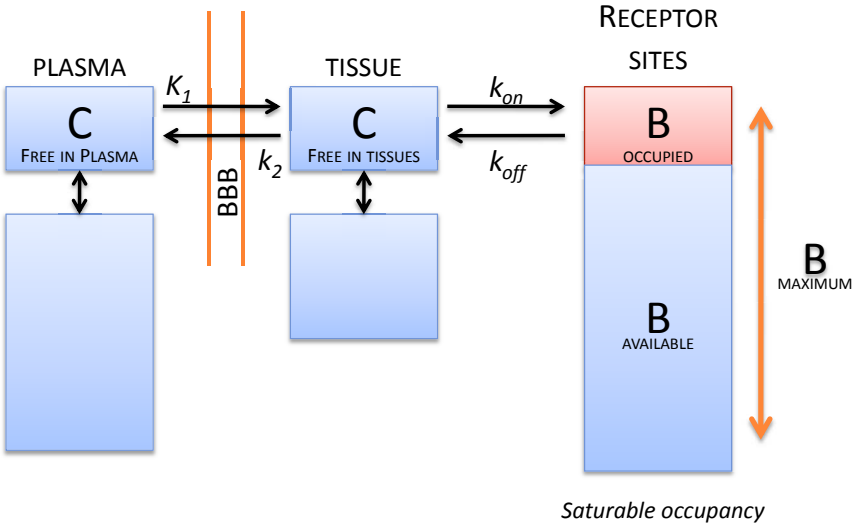

Figure C.2.1: Schematic representation of a general PK-receptor binding model. K_1 and k_2 represent the plasma to brain tissue and brain tissue to plasma constant rates respectively. BBB indicates the blood-brain barrier. k_{on} and k_{off} represent the association and dissociation rates between free brain concentration and receptor.

PET for the PD drug characterization does not usually involve more than 10-12 subjects by limiting the precision and accuracy with which the drug PK/PD behaviour is characterized. Secondly, in a PET scan several unexpected events (like patient PET scanner intolerances, failures due to the synthesis of the tracer,...) can occur with the effect of delaying or eventually eliminating its execution. In addition, when more than one PET scan has to be performed per patient, there is a minimum amount of time (generally 4 hours) in which no other PET scans can be performed. As consequence of these two factors, the time location of PET scans practicable in reality cannot cover all the possible combinations definable in theory.
In common practice the scan times are define targeting the maximal plasma concentration and lower concentration (usually 24 hours post dose) of drug PK[17]. For what concerns the drug dose, it is generally selected on the basis of pre-clinical information and early human data. However, since this information might be misleading (especially when new drug compound are evaluated), the experimental design can be corrected during the analysis of the patients. Also for this case, however, no particular criteria about the convenience to change or not the experimental design is available. The decision is subordinated to the experience of the experiment management committee.

C.2.3 Optimal Designs in PET occupancy study

Optimal designs are a class of experimental designs that are defined to estimate the parameters of interest with the best precision and accuracy in term of bias and variance. This result is achievable because the optimal designs are based on minimizing the estimated parameters' covariance matrix (Ψ_θ). Ψ_θ is usually unknown but, from the Cramer-Rao theorem[18], can be lower bounded by the inverse of the Fisher Information Matrix (FIM)[19]. This can be explicitly indicated in

$$\Psi_\theta \geq (FIM)^{-1} \tag{C.3}$$

There are different kinds of optimal design algorithms depending on the system model and the statistical criterion for the information maximization. The latter refers to the way in which Ψ_θ is managed, because, from a strictly mathematical point of view, the minimization of a matrix has no meaning. What is usually done is to translate the information contained in the FIM into a scalar number function of experimental parameters that becomes the objective function to minimize. There are several ways

of doing this transformation, thus defining different types of optimal designs. The D-optimality[20,21] is the most common approach and it is based on the determinant of the inverse of FIM:

$$x_D = \arg_x max(|FIM|) \tag{C.4}$$

where x_D represents the optimized setting for the experimental variables. Since the FIM definition is based on the model of the system, any eventually error in the definition of the system model as well as on the identification of its parameters can heavily condition the performance of the design. The last one is a severe limitation for the application of the methodology. It is clear that before the execution of the experiments there is quite a lot of uncertainty on parameter estimates; however we can at least guess (based on *a priori* knowledge of the system) its value. The less accurate the assumption on the true values of parameter model is, the worse the performance of the optimized experiment are. To extend the use of these methods, different strategies have been successfully proposed[22], although with time consuming solutions[23]. Alternatively one possible strategy is to utilize adaptive approaches in which the prior is updated step by step based on experiment results. This solution represents the fundamental of AOD.
As for fixed optimal design algorithms, several solutions for AOD implementation are available. The one we considered in this work, refers to that proposed by Zamuner and colleagues (2010) and already applied in PET occupancy studies. A schematic representation of this procedure is reported in Figure C.2.2. The particularity of this AOD solution is related to the way in which FIM is updated for each step of the study, by combining the information of performed experiments with that to be preformed. For further information the interesting reader is referred to the original work[16]. As for all the AOD methods, also for this approach the larger the number of the experiments is the more accurate the final estimates are. However, from the practical point of view of PET-RO studies, a trade-off between study duration, number of subjects, and trial costs has to reach, limiting the total number of performable scans. Thus the convenience of applying AOD respect to fixed approaches like D-optimality, needs to be carefully analyzed by considering all the operative conditions in which PET-RO studies are performed.

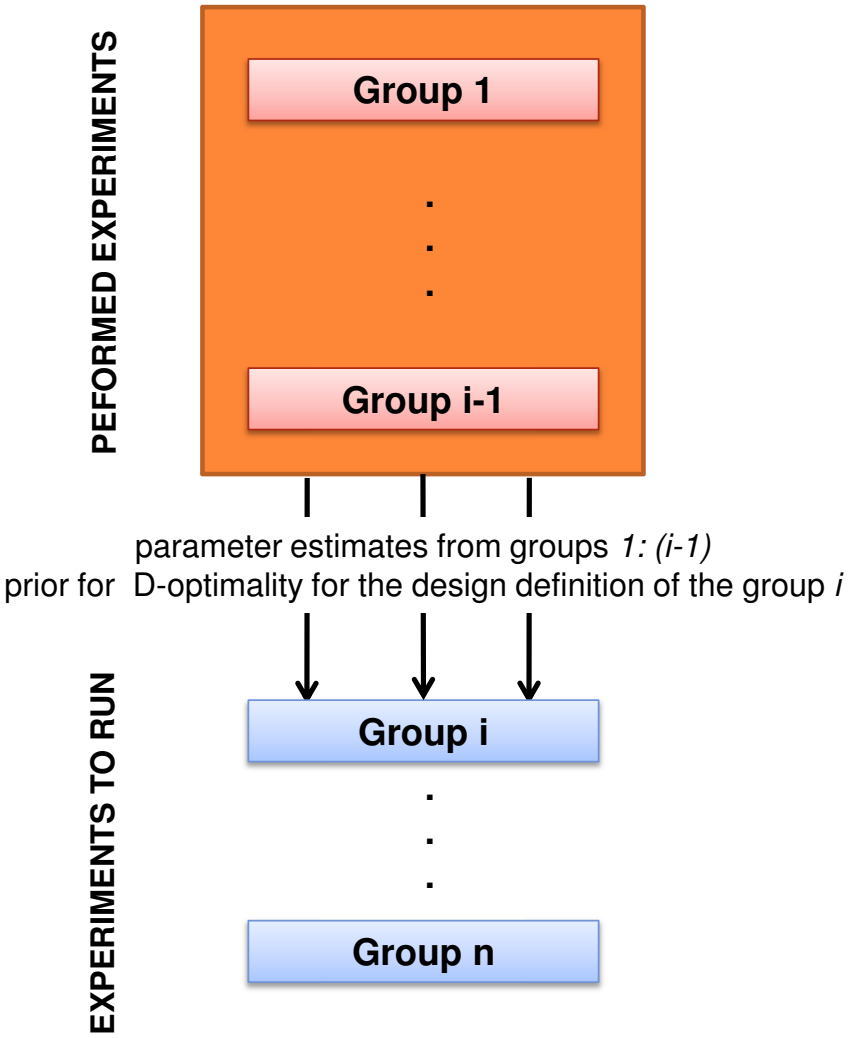

Figure C.2.2: AOD sequentially estimation and optimization algorithm. Adaptive optimization is based on the idea that the estimates derived from the sub population composed by groups $1:(i-1)$ becomes the prior for the design optimization of i^{th} group. This operation is repeated until the last group n. The parameter prior for the first group is derived from *a priori* assumptions about the system of interest. In vitro experiments or literature values might be used to face this step.

C.2.4 Simulation study

In order to compare the performance of AOD with those provided by D-optimality a simulation study was performed, with particular attention to the timing of PET scans and the effects of grouping. Simulated data were generating according the PK-BP model presented in Eq. C.2. One hundred populations each with 12 subjects were considered. Three simulated compounds with different brain affinities (low, medium and high) were tested, with K_d (Eq. C.2) equal to 15, 2.5 and 0.25 respectively. The dose level was held constant for all the simulations in order to evaluate the impact of different drug affinities without any compensation with the dose. A complete description of the simulated parameters is summarized in Table C.2.1. An exponential distribution model was assumed for the inter-subject variability (CV=30%) of BP_0, k_{on} and k_{off} and a proportional error model was assumed for the residual variability (CV=10%). A two-compartmental model with first order absorption was applied for the plasma concentration kinetics and individual PK parameters were assumed known in the PK-BP model. The population PK model was parametrized as clearance (CL=34.9 L/h), volume of distribution at steady state (Vss=1200 L), inter-compartmental clearance (Q=21.7 L/h), fractional central volume (FVC=0.671) and rate of absorption (ka=0.605 hrs-1). Exponential distribution for inter-subject variability was also used for the PK parameters (CVs assumed in the 30 to 45% range). A representation of simulated parameter distribution (for the medium affinity compound) and simulated BP time course is reported in Figure C.2.3.
Simulated experimental designs were chosen according to adaptive (AOD), non-adaptive (D-optimal) optimal designs and non optimized (Empirical) designs by using different levels of parameter misspecifications respect to the true simulated values (range: [-300%;+300%]). Only two time points were assumed per subject, chosen in a time window of 0-36 hours (minimum distance 4 hours). Data simulation and parameter estimation were performed by using AODware, an in-house Matlab-based software package. In AODware data quantification uses NONMEM[24] while experimental optimization was performed by using POPED[25].
Result comparison was done by computing the bias (Bias%) as well as the mean square error (RMSE%) of the estimated parameter values with respect to the true parameter values. With these two performance indexes result precision and accuracy was evaluated.

Table C.2.1: Simulated parameter values

Drug Affinity	k_{on} [hours^{-1}]	k_{off} [hours^{-1}]	K_d [unitless]	Affinity $(1/K_D)$ [unitless]	BP_0 [unitless]
Low	0.044	0.66	15.1	0.07	3
Medium	0.088	0.22	2.5	0.4	3
High	0.440	0.11	0.25	4.0	3

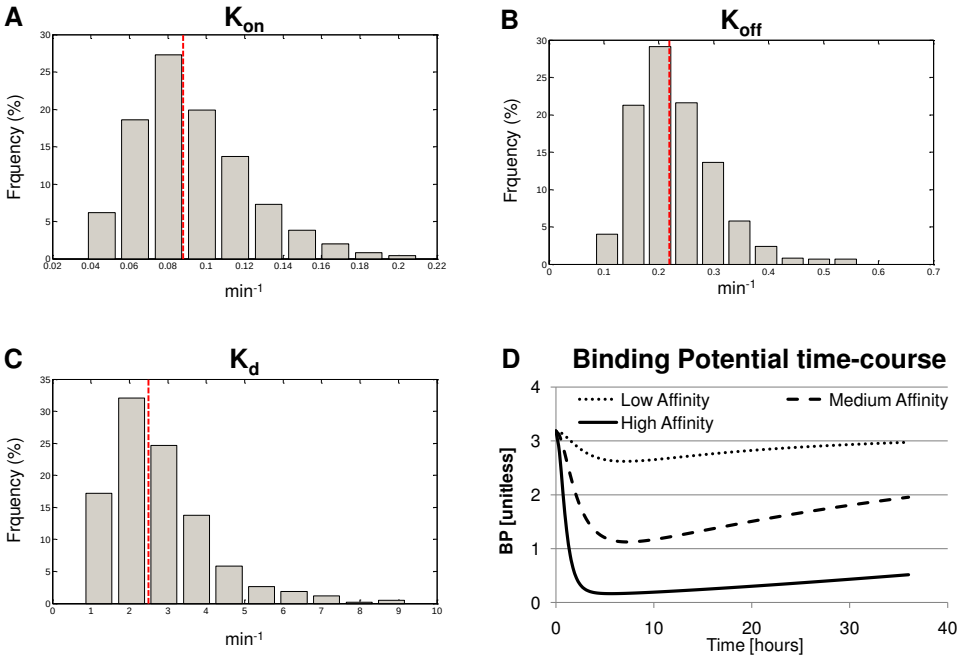

Figure C.2.3: Simulated parameter distributions and BP time-course. Panel A, B and C represent respectively the inter-subject variability for k_{on}, k_{off} and K_d parameters, simulated for the medium affinity drug. For each parameter dashed line represents the population mean. In Panel D the population means of binding potential time courses are reported for all the three tested scenarios: dotted line represents low affinity drug; dashed line represents the medium affinity drug; solid line represents the high affinity drug.

C.3 RESULTS

C.3.1 THE EFFECTS OF GROUPING

Figure C.3.1 shows the performance in the estimation of K_d parameter of D-optimality compared to a fixed empirical design approach. Different number of cohorts within the population were considered. No misspecification was applied for the computation of D-optimal designs. From the analysis of the performance indexes two results are particularly evident: 1) D-optimality provides better performances than the empirical approaches using both unique or multi-groups designs; 2) with the increasing of affinity, there is an evident increasing of precision and accuracy with which the parameters are estimated. This last condition was verified for all the tested designs.
Similar results were obtained for BP_0, while a much higher sensitivity was found for k_{on} and k_{off} parameters. For these parameters the application of D-optimality produced a mean reduction of 500% for the bias% and 300% for RMSE% respected to the empirical designs. This result was achieved for considering both unique and multi-group studies. As for K_d, the results of BP-model parameters (BP_0, k_{on} and k_{off}) for the high affinity compound were superior to those for the low one.
The effects of grouping with D-optimality was almost negligible, with a mean bias% difference lower than -1% for all the parameters. This result is in agreement with the features of D-optimality and with the fact that when it is applied without any misspecification on the prior parameters D-optimality returns the best possible design, independently from the population group organization. The empirical-based design, instead, demonstrated to perform generally better in presence of multi-groups design. The mean differences between bias% reported using unique and multi-groups designs were -2% for K_d, -61% for k_{on} and -52% for k_{off}. Similar results were reported for RMSE%. Considering all the tested designs, all the affinity levelsand all the parameters, the application of multi-groups populations led to more precise and accurate results respect to unique-group design. Thus, we decided to consider only populations with multi cohorts, organized in 3 groups with 4 subjects each.

C.3.2 THE EFFECTS OF MISSPECIFICATION

Figure C.3.2A reports the performance of D-optimality in presence of prior parameter misspecification. The sensitivity of the parameters to

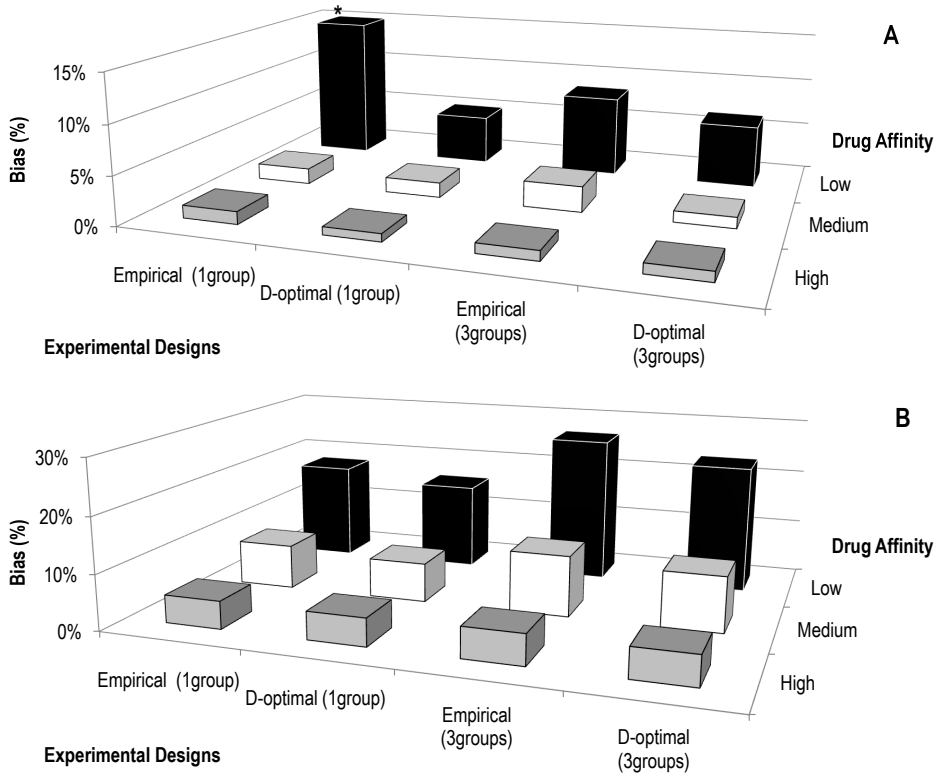

Figure C.3.1: The effects of grouping. Panel A and B represent respectively the K_d bias% and RMSE % obtained for all the simulated levels affinity. Four different scenarios are reported: 1) Empirical design (no optimization applied) for a unique-group population; 2) D-optimal design (no misspecification applied) for a unique-group population; 3) Empirical design (no optimization applied) for a multi-groups population; 4) D-optimal design (no misspecification applied) for a multi-groups population. Empirical design scan timing were fixed to [0,6,24] hours for the unique group population and to [0,6,24], [0,3,12] and [0,8,36] hours for the multi-groups population. In this last case 3 cohorts each one with 4 subjects were considered. * indicates out of scale values.

the experimental design as well as the impact of design on compound with different levels of affinity was found to be consistent with previous results: 1) k_{on} and k_{off} estimates were much more sensitive to the change of time schedule than K_d and BP_0; 2) the higher the affinity the better the performance of the design.
In term of misspecification D-optimality applied to PET occupancy studies is reasonable only when drug affinity misspecification is limited $<$ 100%(absolute values). In other cases, there could be an important bias ($>$50%) in micro (k_{on} and k_{off}) as well as macro (K_d) parameters. In addition it can be see that the presence of positive misspecification on k_{on} and k_{off} parameters does not affect the performance of the estimates as the underestimation of the true parameter values. This fact is consequence of a different location time samples. Figure C.3.2B shows the D-optimal time window define without misspecification and with $\pm 300\%$ of misspecification. When a fast PK/PD drug kinetic is assumed the presence of PET-RO practical constraints (especially the need to wait for at least 4 hours between a PET scan and the next) compensate the effect of misspecification. On the contrary when the behaviour of the compound is assumed lower than the true behaviour (i.e. the case of -300% of misspecification) the timing of PET scans is completely shift from the correct time window.

C.3.3 The effects of AOD

In order to test the capability of AOD to recover from uncorrected parameter assumptions, we tested the methodology to the misspecified scenarios in which D-optimality reported the worst performances (i.e. with $\pm 300\%$ of misspecification on k_{on} and k_{off} parameters), considering the medium affinity simulated compound. The starting point applied to AOD algorithm, i.e. the experimental design for the first cohort of subjects, was exactly the same to the misspecified design indicated by D-optimality. Figure C.3.3A shows the results of AOD on K_d, k_{on} and k_{off} bias%. It is clearly evident how step by step, AOD allows to recover from the presence of misspecification by reducing both bias% and RMSE% compared to the misspecified design used as starting point. Notably, the quality of the final estimates is comparable to those provided by design without misspecification. Compared to the misspecified D-optimal scenario the decreasing of bias% provided by AOD at the third step ranged from 3% (K_d in case of +300% k_{on}/k_{off} misspecification) up to 47% (for k_{on}, in case of -300% k_{on}/k_{off} misspecification). With AOD all the param-

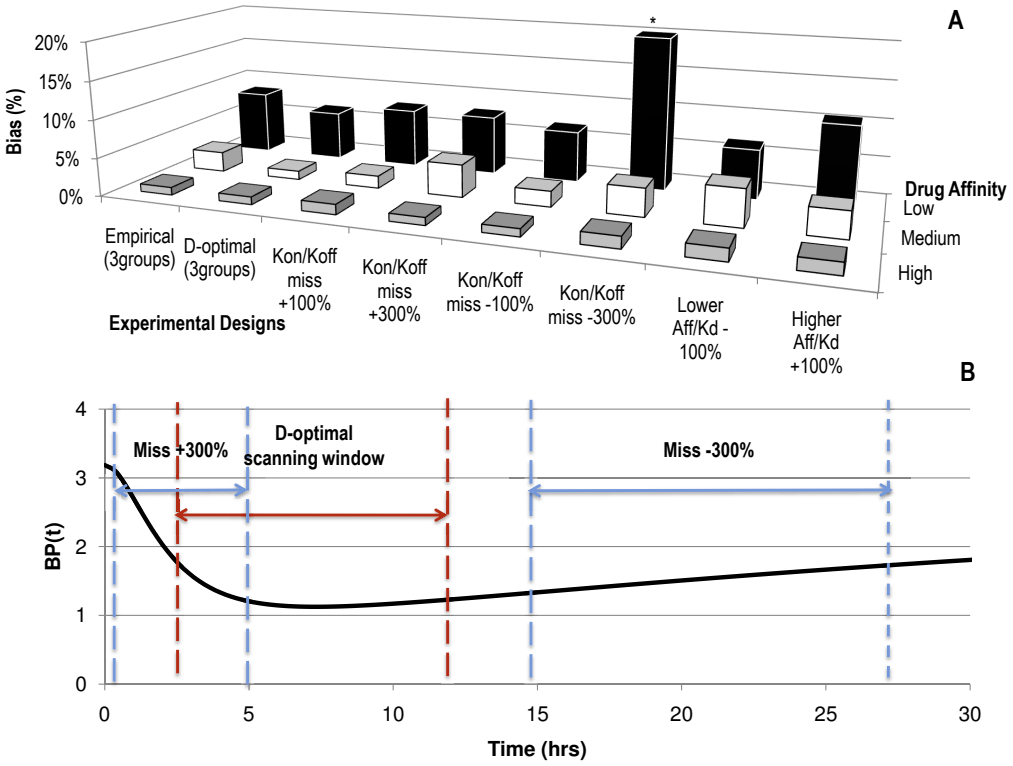

Figure C.3.2: The effects of misspecification on D-optimality. Panel A represents the K_d bias% obtained for all the simulated levels affinity. Results for the empirical design, for the D-optimal design without misspecification and for D-optimal designs applied with different levels of parameter misspecification are reported. All the simulations were performed by using multi-groups populations (3 groups and 4 subjects each). * indicates out of scale value. Panel B reports the scanning time windows obtained by using D-optimality without misspecification (red lines) and with $\pm$300% of parameter misspecification (light blue lines). $BP(t)$ is also shown (solid line).

eters of interest reported bias% $< 10\%$ and RMSE $< 30\%$ in all the tested conditions, while D-optimality produced significantly larger variability (bias% range: 0% - 47%; RMSE% range: 10%-309%). As for the other analysis, the effects of chosen designs are much more evident for k_{on} and k_{off} estimates than for the other parameters of interest (K_d and BP_0).
Figure C.3.3B shows the distribution of AOD scanning times. Results refer to -300% misspecified scenario, but for the correspondent -300% performed very similar. In particular the distributions of AOD scanning times were found to be consistent each other and not influenced by the starting misspecification level. From the figure it emerges that the distributions of PET scan times optimized with AOD are likely aligned with the PET scan times defined with D-optimality without misspecification. This is an important evidence of how well AOD algorithm can correct the initial wrong assumptions about the values of system parameters.

C.4 Considerations

In this work we have presented a simulation study with the final aim to define a some guidelines for a rational management of experimental design variables in PET receptor occupancy studies. In particular, the application of optimal design methods have shown to significantly improve experimental performance, increasing both precision and accuracy with which the PK/PB parameters were estimated. However, results were highly dependent on the choice of optimal design algorithm applied as well as on the level of uncertainty on the true model parameter values. Among all the tested experimental designs, the best performances were achieved using non-adaptive D-optimal method applied without parameter misspecification. Even though, the absence of misspecification is impossible in real practice, the performance obtained in this case represented the best results we should aspired to. At the same time, the worst performances were reported by the same algorithm in presence parameter misspecification. This indicates that the use of D-optimality is adequate only when a complete and sure *a priori* knowledge about the system of interest is available. In light of our results we suggest D-optimality application only when k_{on} and k_{off} misspecification is lower than $< 100\%$.
In agreement with previous studies[16], AOD has shown to valid alternative for PET-RO design definition. In particular, when AOD was applied to high misspecified cases, precision and accuracy of parameters were promptly recovered respect non adaptive methods.

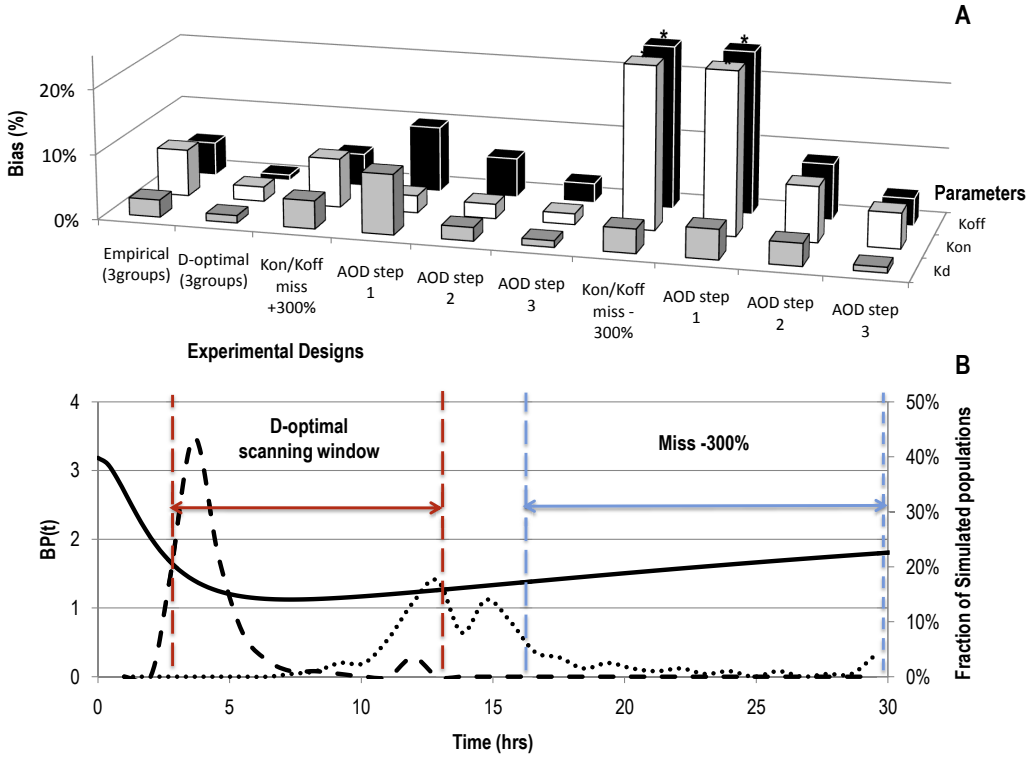

Figure C.3.3: The effects of misspecification on AOD. Panel A represents the K_d, k_{on} and k_{off} bias% obtained for all the simulated drug with medium receptor affinity. Results for the empirical design, for the D-optimal design with and without misspecification and for AOD are reported. All the simulations were performed by using multi-groups populations (3 groups and 4 subjects each). * indicates out of scale value. Panel B reports the scanning time windows obtained by using D-optimality without misspecification (red lines) and -300% of parameter misspecification (light blue lines). Dashed and dotted lines represent the distributions of PET scans obtained by AOD (second and third PET scan respectively). $BP(t)$ is also shown (solid line).

The analysis of simulation results has also offered some interesting considerations about PET-RO experimental criticalities. These can be summarize as follows. Firstly the overestimation of the drug PK/PD behaviour is preferable to an underestimation. In fact, if for the first case the practical constraints of PET-RO experiment allow to compensate for the error, the underestimation of the drug kinetic and dynamic could lead to an unrealistic location of the PET scans respect to the true behaviour of the compound.
Secondly, the organization of the studied population in cohorts as well as the level of drug affinity were found to be two important characteristics to taken into account for the experiment design definition. Combined with the PET scan timing, these aspects were shown to directly affect the precision and the accuracy of parameter estimates. In particular, high-affinity compounds and multi-cohort populations were found to be much more robust to experimental setting changes respect to low affinity drugs and unique-group populations.
Moreover, the sensitivity PK/BP parameters to the experimental design was found to be significantly different: K_d was the most robust parameter (bias% range [1%; 30%]), while k_{on} and k_{off} were much more sensitive to experimental choices (maximum bias% 64% and 50% respectively). This indicates that a correct characterization of the drug affinity is much more easier respect to the identification of binding potential rate constants.
In conclusion, a rational location of the PET scan time, defined according to the characteristics of the tested compounds (kinetics, dynamic and inter-subjects variability) can greatly ameliorate the quality of the estimates in PET receptor occupancy studies. This result can be achieved by organizing the studied population in cohorts with different experimental designs and by applying the available methods for experimental design optimization. In particular adaptive optimization, compared to the fixed approaches, is shown to be robust to the misspecified prior parameter assumptions and therefore it might represent a valid tool to support the user for experimental design definition in PET-RO studies.

C.5 References

[1] R.S. Uppoor, P. Mummaneni, E. Cooper, H.H. Pien, A.G. Sorensen, J. Collins, M.U. Mehta, S.U. Yasuda. The use of imaging in the early development of neuropharmacological drugs: a survey of approved NDAs. Clin Pharmacol Ther. 84(2008), 69-74.

[2] O. Gefvert, M. Bergstrom, B. Langstrom, T. Lundberg, L. Lindstrom, R. Yates. Time course of central nervous dopamine-D2 and 5-HT2 receptor blockade and plasma drug concentrations after discontinuation of quetiapine (Seroquel) in patients with schizophrenia. Psychopharmacology (Berl). 135(1998),119-26.
[3] A.M. Catafau, M.M. Penengo,G. Nucci, S. Bullich, I. Corripio, E. Parellada, C. García-Ribera, R. Gomeni, E. Merlo-Pich. Pharmacokinetics and time-course of D(2) receptor occupancy induced by atypical antipsychotics in stabilized schizophrenic patients..J Psychopharmacol. 22(2008), 882-94.
[4] F. Mentré, A. Mallet, D. Baccar. Optimal design in random-effects regression models. Biometrika. 84(1997), 429-442.
[5] A.C. Hooker, P. Vicini. Simultaneous population optimal design for pharmacokinetic-pharmacodynamic experiments. AAPS J. 7(2005), 759-85.
[6] S. Duffull, F. Mentré, L. Aarons. Optimal Design of a Population Pharmacodynamic Experiment for Ivabradine. Pharm Res. 18 (2001), 83-89.
[7]S. Retout, S. Duffull, F. Mentré. Development and implementation of the population Fisher information matrix for the evaluation of population pharmacokinetic designs. Comput. Meth. Prog. Biomed. 65(2001), 141-51.
[8] A.C. Hooker, M. Foracchia, M. G. Dodds, P. Vicini. An evaluation of population D-optimal designs via pharmacokinetic simulations, Annals Of Biomedical Engineering. 31(2003), 98-111.
[9] E. Walter, L. Pronzato, Optimal experiment design for nonlinear models subject to large prior uncertainties, Am. J. Physiol. 253(1987) R530-R534.
[10] R. R. Sitter. Robust designs for binary data. Biometrics 48(1992), 1145-1155 .
[11] E. M. Landaw. Robust sampling designs for compartmental models under large prior eigenvalue uncertainties. In: Mathematics and Computers in Biomedical Applications, (J.Eisendfeld and C. DeLisi Eds.), Elsevier Science, North Holland, (1985) 181-187.
[11] M. Tod, J.M. Rocchisani. Comparison of ED, EID and API criteria for the robust optimization of sampling times in pharmacokinetics. J. Pharmacokin. Biopharm. 25(1997), 515-537.
[12] K. Ogungbenro, I. Gueorguieva, O. Majid, G. Graham, L. Aarons. Optimal design for multiresponse pharmacokinetic-pharmacodynamic models - dealing with unbalanced designs. J Pharmacokinet Pharmacodyn. 34(2007), 313-31.

[13] K. Ogungbenro, L. Aarons. Optimisation of sampling windows design for population pharmacokinetic experiments. J Pharmacokinet Pharmacodyn. 35(2008), 465-482.
[14] J.J. Di Stefano II, Optimized blood sampling protocols and sequential design of kinetic experiments. Americ. J. Physiol. 240(1981), 259-265.
[15] K. Ogungbenro, A. Dokoumetzidis, L. Aarons. Application of optimal design methodologies in clinical pharmacology experiments. Pharm Stat. 8(2009), 239-52.
[16] S. Zamuner, V.L. Di Iorio, J. Nyberg, R.N. Gunn, V.J. Cunningham, R. Gomeni, A.C. Hooker. Adaptive-optimal design in PET occupancy studies. Clin Pharmacol Ther. 87(2010), 563-571.
[17] D. Mamo, S. Kapur, C.M. Shammi, G. Papatheodorou, S. Mann, F. Therrien, G.A. Remington. PET study of dopamine D2 and serotonin 5-HT2 receptor occupancy in patients with schizophrenia treated with therapeutic doses of ziprasidone. Am J Psychiatry. 161(2004), 818-25.
[18] R. J. Larsen, M. L. Marx. An Introduction to Mathematical Statistics and its Applications, Englewood Cliffs, NJ, Prentice-Hall, 1986.
[19] J.W. Pratt, F. Y. Edgeworth and R. A. Fisher on the Efficiency of Maximum Likelihood Estimation. The Annals of Statistics. 4(1976), 501-514.
[20] V.V. Fedorov, Theory of Optimal Experiment, Academic Press, New York, 1972.
[21] D. Z. DArgenio. Incorporating prior parameter uncertainty in the design of sampling schedules for pharmacokinetic parameter estimation experiments. Math. Biosci. 99(1990), 105-118.
[22] M.G. Dodds, A.C. Hooker, P. Vicini. Robust population pharmacokinetic experiment design. J Pharmacokinet Pharmacodyn. 32(2005), 33-64.
[23] E.N. Jonsson, J.R. Wade, M.O. Karlsson. Comparison of some practical sampling strategies for population pharmacokinetic studies. J. Pharmacokinet. Biopharm. 24 (1996), 245-263.
[24] S. Beal, L.B. Sheiner NONMEM User Guides. (Nonmem Project Group 1998).
[25] M. Foracchia, A. Hooker, P. Vicini, A. Ruggeri. POPED, a software for optimal experiment design in population kinetics. Comput Methods Programs Biomed, 74, 29-46 (2004).

C.6 Project Partnership

This research was conducted in collaboration with prof. Roger Gunn (Imanova Limited, London, UK; Imperial College, London, UK) and Ste-

fano Zamuner (GSK, London, UK). Portions of this work were presented at XXIth PAGE meeting (Venice, 2012).

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor Prof. Alessandra Bertoldo for the continuous support of my Ph.D study and research, for her patience, motivation, enthusiasm and all the possibilities that she offered me during these years we have been working together: I could not have imagined having a better mentor.

Besides my advisor, I would like to thank "my colleague in PET" Dr. Gaia Rizzo, for thousands of reasons that she knows better than anybody else.

My sincere thanks also goes to all the Functional and Anatomical Imaging Research group members: Letizia Squarcina, Denis Peruzzo, Marco Castellaro, Elisa Veronese and Enrico Grisan. We had great time together.

Thank you to everyone that has contributed to the realization of this work.